

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

Searcher: JOHN DANTZMAN

Searcher Phone #: 308 4488

Searcher Location: _____

Date Searcher Picked Up: 5-1-01

Date Completed: 5-2-01

Searcher Prep & Review Time: 10

Clerical Prep Time: _____

Online Time: 10

Type of Search

NA Sequence (#) _____

AA Sequence (#) 5

Structure (#) _____

Bibliographic _____

Litigation _____

Fulltext _____

Patent Family _____

Other _____

Vendors and cost where applicable

STN _____

Dialog _____

Questel/Orbit _____

Dr.Link _____

Lexis/Nexis _____

Sequence Systems ABSS02

WWW/Internet _____

Other (specify) _____

THIS PAGE BLANK (USPTO)

STIC-Biotech/ChemLib

41408

From: Hutzell, Paula
Sent: Monday, April 30, 2001 4:51 PM
To: STIC-Biotech/ChemLib; Graser, Jennifer
Subject: FW: rush search

Importance: High

please rush

-----Original Message-----

From: Graser, Jennifer
Sent: Monday, April 30, 2001 3:57 PM
To: Hutzell, Paula
Subject: rush search
Importance: High

Hi Paula,

Could you please authorize the following rush search for an Election?

Thanks,
Jennifer

STIC:

Please search SEQ ID NOs: 1-5 from 09/142,970 in pending and commercial databases.

Thanks,
Jennifer Graser
CM1 7E09 (mailbox 7E12)
Art Unit 1645
308-1742

Declarat
say s filed
as PCT
EP 98/00294 11/29/97
at top
then below say

1/20/98

prior
1/21/97

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

Seg 1

OM protein - protein search, using sw model

Run on: May 1, 2001, 14:48:53 ; Search time 37.5 seconds
(without alignments)
53.278 Million cell updates/sec

Title: US-09-142-970-1

Perfect score: 558
Sequence: 1 LYKNNRYALKSGSVNAP.....NRFLLTGCTNLNGKISVTG 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapect 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6C.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/6D.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	485	86.9	1507	6	5268270-2
2	329.5	59.1	1541	5	PCT-US95-10661A-3
3	322.5	57.8	1702	5	PCT-US95-10661A-5
4	320.5	57.4	1545	5	PCT-US95-10661A-4
5	311.5	55.8	1848	5	PCT-US95-10661A-6
6	114	20.4	1394	5	PCT-US95-10661A-2
7	75	13.4	1377	2	US-08-793-824-2
8	73.5	13.2	631	1	US-08-487-890A-115
9	73.5	13.2	631	2	US-08-478-435-115
10	73.5	13.2	631	2	US-08-337-483-115
11	73.5	13.2	631	2	US-08-478-373-115
12	73.5	13.2	631	3	US-08-474-671-115
13	73.5	13.2	631	3	US-08-483-577A-115
14	72	12.9	430	2	US-08-945-848-8
15	71	12.7	806	1	US-07-980-528-2
16	67.5	12.1	631	1	US-08-487-890A-111
17	67.5	12.1	631	2	US-08-478-435-111
18	67.5	12.1	631	2	US-08-337-483-111
19	67.5	12.1	631	2	US-08-478-373-111
20	67.5	12.1	631	3	US-08-474-671-111
21	67.5	12.1	631	3	US-08-483-577A-111
22	67.5	12.1	890	2	US-08-483-101-14
23	66	11.8	434	2	US-08-795-475-3
24	66	11.8	648	1	US-08-487-890A-109
25	66	11.8	648	2	US-08-478-435-109
26	66	11.8	648	2	US-08-337-483-109
27	66	11.8	648	2	US-08-478-373-109

28	66	11.8	648	3	US-08-474-671-109	Sequence 109, App
29	66	11.8	648	3	US-08-483-577A-109	Sequence 109, App
30	64.5	11.6	275	5	US-08-900-565-4	Sequence 4, Appl
31	64.5	11.6	453	6	5510466-4	Patent No. 5510466
32	64.5	11.6	1618	1	US-07-853-913-4	Sequence 6, Appl
33	64	11.5	394	3	US-08-673-814-6	Sequence 6, Appl
34	64	11.5	1338	2	US-08-728-470-9	Sequence 9, Appl
35	64	11.5	1599	2	US-08-617-697-9	Sequence 9, Appl
36	63	11.3	708	3	US-08-613-009A-22	Sequence 22, Appl
37	63	11.3	1529	2	US-08-728-470-10	Sequence 10, Appl
38	63	11.3	1600	2	US-08-617-697-10	Sequence 10, Appl
39	62.5	11.2	671	2	US-08-737-716-13	Sequence 13, Appl
40	62	11.1	305	2	US-08-795-475-1	Sequence 1, Appl
41	62	11.1	630	1	US-08-487-890A-113	Sequence 113, App
42	62	11.1	630	2	US-08-478-435-113	Sequence 113, App
43	62	11.1	630	2	US-08-337-483-113	Sequence 113, App
44	62	11.1	630	2	US-08-478-373-113	Sequence 113, App
45	62	11.1	630	3	US-08-474-671-113	Sequence 113, App

ALIGNMENTS

RESULT 1
5268270-2
Patent No. 5268270
APPLICANT: Meyer, Thomas F.; Halter, Roman; Pohlner, Johannes
TITLE OF INVENTION: PROCESS FOR PRODUCING PROTEINS USING GRAM
NEGATIVE HOST CELLS
NUMBER OF SEQUENCES: 6
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/171,872
FILING DATE: 01-JUL-1987
SEQ ID NO: 2
LENGTH: 1507
5268270-2

Query Match 86.9%; Score 485; DB 6; Length 1507;
Best Local Similarly 86.5%; Pred. No. 4.8e-50;
Matches 90; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 LYKNNRYALKSGSVNAPPENGOTENNDIIMSGTOEAKNNHKNRISFSG 60
DB 559 LYKNNRYALKSGGRNAPPENGVAENNDIIMSGTOEAKNNHKNRIGDEGC 618
QY 61 FFGENGKGHGNALNPNCKSAONRFLTGCTNLNGKISVTG 104
DB 619 FFDENGKGHGNALNPNCKSAONRFLTGCTNLNGKISVTG 662

RESULT 2
PCT-US95-10661A-3
Sequence 3, Application PC/TUS9510661A
GENERAL INFORMATION:
APPLICANT: Washington University, et al.
TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10661A
FILING DATE: 16-AUG-1995

OPERATING SYSTEM: PC-DOS/MS-DO
SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10661A
FILING DATE: 16-AUG-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,791
FILING DATE: 25-AUG-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Treacartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: RP-55941/RFT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-10661A-2

Query Match 20.4%; Score 114; DB 5; Length 1394;
Best Local Similarity 38.3%, Pred. No. 3,6e-05;

Matches 23; Conservative 13; Mismatches 24; Indels 0; Gaps 0.

OY 45 NAMHNKNNRISGSGFGENGKGHNCALNENGSXONRFLLTGTGNLNGKTSVYG 104
Dn 551 NNINKLDYRKRIAVANGWGETDKNKHNGRLNLHYRPETEDPTLLSGGTNLKGDITQIKG 610

RESULT 7
US-08-793-824--2
Sequence 2, Application US/08793824
Patent No. 5981838
GENERAL INFORMATION:
APPLICANT: Simpson, Christine Lynn
APPlicant: Giffard, Philip Morrison
Applicant: Jacques, Nicholas Anthony
TITLE OF INVENTION: Genetic Manipulation of Plants to
Title of Invention: Increase Stored Carbohydrates
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDressee: Griffith Hack & Co
STREET: Level 8, 168 Walker Street
CITY: No. 5981838th Sydney
STATE: New South Wales
COUNTRY: Australia
ZIP: 2060
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DO
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,824
FILING DATE:
CLASSIFICATION: 800
Prior Application Data:
Application Number: AU PM7643
Filing Date: 24-Aug-1994
Telecommunication Information:
Telephone: 61 2 9957 5944
Telefax: 61 2 957 6288
Telex: 26547
Information For Seq Id No: 2:
Sequence Characteristics:
Length: 1577 amino acids

TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: not relevant
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Streptococcus salivarius
US-08-793-824-2

Query Match 13.4%; Score 75; DB 2; Length 1577;
Best Local Similarity 23.5%; Pred No. 2.2; Mismatches 43; Indels 30; Gaps 4;
Matches 27; Conservative 15; Mismatches 43; Indels 30; Gaps 4;

QY 5 NYRYALKSGSVNAPMPENQOTENNMDILMGSTOEAKKNAMNKNRISG---FSG 60
DB 1374 NGRYFLDGDGSGEL---APSRFVTENNKKYVYDG-----NGKLVGCAVINGNNHYFNN 1423

OY 61 FTGENGKGNALNINFGKSAQNRFL-----LTGTINLNGKI 99
DB 1424 DYSQVKGAMNGRYDDGSGAIVSNQFIQIANOMAYLNDGCHKYGLONINNKV 1478

RESULT 8

US-08-487-890A-115
Sequence 115, Application US/08487890A
Patent No. 5708149

GENERAL INFORMATION:

APPLICANT: Loomore, Sheena
APPLICANT: Harkness, Robin
APPLICANT: Schryvers, Anthony
APPLICANT: Chong, Pele
APPLICANT: Gray-Owen, Scott
APPLICANT: Yang, Yan-ping
APPLICANT: Mordin, Andrew
APPLICANT: Klein, Michel
TITLE OF INVENTION: Transferrin Receptor Genes
NUMBER OF SEQUENCES: 147
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,890A
FILING DATE: 07-JUN-1993
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/175,116
FILING DATE: 29-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,968
FILING DATE: 08-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:
LENGTH: 631 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-487-890A-115

Query Match 13.2%; Score 73.5; DB 1; Length 631;
Best Local Similarity 27.4%; Pred No. 1; Mismatches 50; Indels 17; Gaps 5;
Matches 29; Conservative 10; Mismatches 50; Indels 17; Gaps 5;

QY 2 YKKNRYALKSGSVNAPMPENQOTENNMDILMGSTOEAKKNAMNKNRISG---FSG 58
DB 153 FSGYGYAVYFGKQATATLPVNGEATYKGTWSEITATFERGKNSLNNRGQVSRRSAT 212

OY 59 SGFFGENG-KGNALNINFGKSAQNRFLTG-----GTINL 96
DB 213 PDDIDLENDAGLTSEFTVNFQTK-----LTGEPYNERETINL 252

RESULT 9

US-08-478-435-115
Sequence 115, Application US/08478435
Patent No. 5922323

GENERAL INFORMATION:

APPLICANT: Loomore, Sheena
APPLICANT: Harkness, Robin
APPLICANT: Schryvers, Anthony
APPLICANT: Chong, Pele
APPLICANT: Gray-Owen, Scott
APPLICANT: Yang, Yan-ping
APPLICANT: Mordin, Andrew
APPLICANT: Klein, Michel
TITLE OF INVENTION: Transferrin Receptor Genes
NUMBER OF SEQUENCES: 147
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,435
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/337,483
FILING DATE: 08-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175,116
FILING DATE: 29-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/148,968
FILING DATE: 08-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-462 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:
LENGTH: 631 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-478-435-115

Query Match

13.2%; Score 73.5; DB 2; Length 631;

Best Local Similarity 27.4%; Pred. No. 1;
Matches 29; Conservative 10; Mismatches 50; Indels 17; Gaps 5;

OY 2 YKKNRYALKSGSVNAPBENGQ--TENNDWILMGSTOEAKKNMHNKN--QRTSGF 58
Db 153 FYSGYGYATYFGKQTATTLVPNGEATYKGTWSEFTATETGKNYSLENNNGOAYSRSRSAT 212
OY 59 SGFFGEENG-KGHNGALNLFNGKSAONRFLTG-----GTNLN 96
Db 213 PGDIDLENGDAGLTSEFTVNFGTK-----LTGEPYNNERTN 252

RESULT 10

US-08-337-483-115
; Sequence 115, Application US/08337483
; Patent No. 5922562
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena
; APPLICANT: Harkness, Robin
; APPLICANT: Schryvers, Anthony
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Mirdin, Andrew
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Transferrin Receptor Genes
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/337,483
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-410 MIS:7b
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 631 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-337-483-115

Query Match 13.2%; Score 73.5; DB 2; Length 631;
Best Local Similarity 27.4%; Pred. No. 1;
Matches 29; Conservative 10; Mismatches 50; Indels 17; Gaps 5;

OY 2 YKKNRYALKSGSVNAPBENGQ--TENNDWILMGSTOEAKKNMHNKN--QRTSGF 58
Db 153 FYSGYGYATYFGKQTATTLVPNGEATYKGTWSEFTATETGKNYSLENNNGOAYSRSRSAT 212
OY 59 SGFFGEENG-KGHNGALNLFNGKSAONRFLTG-----GTNLN 96
Db 213 PGDIDLENGDAGLTSEFTVNFGTK-----LTGEPYNNERTN 252

RESULT 11

US-08-478-373-115
; Sequence 115, Application US/08478373
; Patent No. 5922841
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena
; APPLICANT: Harkness, Robin
; APPLICANT: Schryvers, Anthony
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Mirdin, Andrew
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Transferrin Receptor Genes
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,373
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,483
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,968
; FILING DATE: 08-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-463 MIS:v9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 631 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-478-373-115

Query Match 13.2%; Score 73.5; DB 2; Length 631;
Best Local Similarity 27.4%; Pred. No. 1;
Matches 29; Conservative 10; Mismatches 50; Indels 17; Gaps 5;

OY 2 YKKNRYALKSGSVNAPBENGQ--TENNDWILMGSTOEAKKNMHNKN--QRTSGF 58
Db 153 FYSGYGYATYFGKQTATTLVPNGEATYKGTWSEFTATETGKNYSLENNNGOAYSRSRSAT 212
OY 59 SGFFGEENG-KGHNGALNLFNGKSAONRFLTG-----GTNLN 96
Db 213 PGDIDLENGDAGLTSEFTVNFGTK-----LTGEPYNNERTN 252

RESULT 12
US-08-474-671-115
; Sequence 115, Application US/08474671
; Patent No. 6008326

GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena
APPLICANT: Harkness, Robln
APPLICANT: Schryvers, Anthony
APPLICANT: Chong, Pele
APPLICANT: Gray-Owen, Scott
APPLICANT: Yang, Yan-Ping
APPLICANT: Murdln, Andrew
TITLE OF INVENTION: Transferrln Receptor Genes
NUMBER OF SEQUENCES: 147
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slm & Mcburney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,671
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/337,483
FILING DATE: 08-NOV-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175,116
FILING DATE: 29-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,968
FILING DATE: 08-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-465 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:
LENGTH: 631 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-474-671-115

Query Match 13.2%; Score 73.5; DB 3; Length 631;
Best Local Similarity 27.4%; Pred. No. 1;
Matches 29; Conservative 10; Mismatches 50; Indels 17; Gaps 5;

Oy 2 YKKNRYVALSGSVNAMPENGO-TENNNDWILMGSTOEAKKNAHKN--ORISGF 58
Db 153 FYSGYGYAIFYFGKGTATTLTPVNGEATYKGTWSTTATERCKNYSLFNNRGQAYSRRSAT 212
Oy 59 SGFGEENG-KGHNGALNLFNGKSAQNRFLTG-----GTNIN 96
Db 213 PGDIDLENGDAGLTSEFTVNFGTKK-----LTGEPYNERETNIN 252

RESULT 13
US-08-483-577A-115
Sequence 115, Application US/08483577A
Patent No. 6015688
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena
APPLICANT: Harkness, Robln

APPLICANT: Schryvers, Anthony
APPLICANT: Chong, Pele
APPLICANT: Gray-Owen, Scott
APPLICANT: Yang, Yan-Ping
APPLICANT: Murdln, Andrew
TITLE OF INVENTION: Transferrln Receptor Genes
NUMBER OF SEQUENCES: 160
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slm & Mcburney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/483,577A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/337,483
FILING DATE: 08-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175,116
FILING DATE: 29-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,968
FILING DATE: 08-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-511
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 115:
SEQUENCE CHARACTERISTICS:
LENGTH: 631 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-483-577A-115

Query Match 13.2%; Score 73.5; DB 3; Length 631;
Best Local Similarity 27.4%; Pred. No. 1;
Matches 29; Conservative 10; Mismatches 50; Indels 17; Gaps 5;

Oy 2 YKKNRYVALSGSVNAMPENGO-TENNNDWILMGSTOEAKKNAHKN--ORISGF 58
Db 153 FYSGYGYAIFYFGKGTATTLTPVNGEATYKGTWSTTATERCKNYSLFNNRGQAYSRRSAT 212
Oy 59 SGFGEENG-KGHNGALNLFNGKSAQNRFLTG-----GTNIN 96
Db 213 PGDIDLENGDAGLTSEFTVNFGTKK-----LTGEPYNERETNIN 252

RESULT 14
US-08-945-848-8
Sequence 8, Application US/08945848
Patent No. 5868772
GENERAL INFORMATION:
APPLICANT: MATSUSHIRO, Aizo
TITLE OF INVENTION: PEARL PROTEIN(NACREIN) AND PROCESS FOR
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 14:50:33 ; Search time 43.98 Seconds

(without alignments)
162,510 Million cell updates/sec

Title: US-09-142-970-1

Perfect score: 558

Sequence: 1 LYKKNRYRYALKSGGSVNAP.....NRFLLTGGTNLNGKISVTGQ 104

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

PIR-67:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	558	100.0	1561	2 S61314	IGA-specific metal
2	558	100.0	1773	2 A81937	IGA-specific metal
3	554	99.3	1815	2 C81169	IGA-specific metal
4	488	87.5	1532	2 A26039	IGA-specific metal
5	329.5	59.1	1541	2 A37023	IGA-specific metal
6	328.5	58.9	1694	2 H64106	IGA-specific metal
7	328.5	58.9	1702	2 A41859	IGA-specific metal
8	320.5	57.4	1545	2 B41859	IGA-specific metal
9	317.5	56.9	1849	2 C41859	IGA-specific metal
10	128	22.9	1431	2 A81018	serine-type peptid
11	115.5	20.7	709	2 C64057	IGA-specific metal
12	115	20.6	1449	2 B81963	IGA-specific metal
13	114	20.4	1394	2 S60762	IGA-specific metal
14	111	19.9	1457	2 D81019	adhesion and penet
15	84.5	15.1	243	2 T24981	hypothetical prote
16	82.5	14.8	451	2 A23535	clustered asparagi
17	80.5	14.4	642	2 D81401	probable flagellar
18	80	14.3	629	2 B75330	probable ribosomal
19	79	14.2	2529	2 B64635	toxin-like outer m
20	78	14.0	568	2 JC7210	mollican shell ma
21	78	14.0	2399	2 H71879	toxin-like outer m
22	77.5	13.9	419	2 T40014	probable vesicular
23	76.5	13.7	1650	2 T18444	hypothetical prote
24	76	13.6	321	2 B37050	porin precursor -
25	75.5	13.5	484	2 G70846	hypothetical glyci
26	75.5	13.5	954	1 S20907	endo-1,4-beta-xyla
27	75	13.4	697	2 T16306	glucosyltransferas
28	75	13.4	1577	2 T30858	ARPI protein - yea
29	74.5	13.4	719	2 S61046	

30	74.5	13.4	1596	2 A33106	neurogenic locus m
31	73.5	13.2	631	2 S70910	transferrin-bindin
32	72.5	13.1	1449	2 T30857	glucosyltransferas
33	72.5	13.0	125	2 T24982	hypothetical prote
34	72.5	13.0	234	2 S14469	asparagine-rich pr
35	72	12.9	402	2 T13614	N-acetyltransferas
36	72	12.9	712	2 E81196	transferrin-bindin
37	72	12.9	1367	2 T18466	hypothetical prote
38	72	12.9	1356	2 S76781	glutamate synthase
39	71.5	12.8	515	2 S20493	endoglucanase - C1
40	71.5	12.8	698	2 JH0162	No-on-transient A
41	71.5	12.8	700	2 H71607	No-on-transient A
42	71.5	12.8	764	2 T23076	hypothetical prote
43	71	12.7	163	2 T23076	hypothetical prote
44	71	12.7	461	2 T10265	arabinogalactan-pr
45	71	12.7	537	2 A23770	asparagine-rich pr

ALIGNMENTS

RESULT 1
S61314 IGA-specific metalloendopeptidase (EC 3.4.24.13) homolog - Neisseria meningitidis
C:Species: Neisseria meningitidis
A:Variety: HF13
C>Date: 20-Jul-1996 #sequence-revision 13-Mar-1997 #text-change 08-Dec-2000
C:Accession: S61314
R:Lomholt, H.; Poulsen, K.; Mogens, K.
Mol. Microbiol. 15, 495-506, 1995
A>Title: Comparative characterization of the iga gene encoding Iga1 protease in Neiss
A:Reference number: S61314; MUID:95302961
A:Accession: S61314
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1561 <LOM>
A:Cross-references: EMBL:X82474; NID:G732873; PIDN:CA57857.1; PID:G732874
C:Superfamily: IGA-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase

Query Match 100.0%; Score 558; DB 2; Length 1561;
Best Local Similarity 100.0%; Pred. No. 1.4e-48;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LYKKNRYRYALKSGGSVNAPDPENGQTEENDWIMSGTQEAKKANNNHKNRISGFSG 60
DB 584 LYKKNRYRYALKSGGSVNAPDPENGQTEENDWIMSGTQEAKKANNNHKNRISGFSG 643
OY 61 PFGEENGKGNHGNALNPNKSAQNRFLNGTGLNKRISVTGQ 104
DB 644 PFGEENGKGNHGNALNPNKSAQNRFLNGTGLNKRISVTGQ 687
RESULT 2
A81937 IGA-specific metalloendopeptidase (EC 3.4.24.13) NMA0905 [imported] - Neisseria menin
N:Alternate names: Iga1 proteinase, Iga1 proteinase (EC 3.4.21.7) [misnomer]; Immunogl
C:Species: Neisseria meningitidis
A:Variety: group A strain 22491; strain HF117; strain HF159; strain SM1027
C>Date: 05-May-2000 #sequence-revision 05-May-2000 #text-change 08-Dec-2000
C:Accession: A81937; S61317; S61318; S61321
R:Perkhill, J.; Achman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mc
Hollroyd, S.; Jørgensen, K.; Leather, S.; Mungall, K.; Quail, M.A.; Rajandhe
Nature 404, 502-506, 2000
A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491
A:Reference number: A81775; MUID:20222556
A:Accession: A81937
A:Molecule type: DNA
A:Residues: 1-1773 <PAR>
A:Cross-references: GB:AL162754; GB:AL157959; NID:G7379424; PIDN:CAB84182.1; PID:G737
A:Experimental source: serogroup A, strain 22491
R:Lomholt, H.; Poulsen, K.; Mogens, K.

Mol. Microbiol. 15, 495-506, 1995
 A>Title: Comparative characterization of the iga gene encoding Iga1 protease in *Neisseria*
 A:Reference number: S61314; MUID:95302961
 A:Accession: S61317
 A>Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 53-548 <LOM>
 A:Cross-references: EMBL:X82470; NID:g732854; PIDN:CA57853.1; PID:g732855
 A:Experimental source: strain HF117
 A:Accession: S61318
 A>Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 53-548 <LOM>
 A:Cross-references: EMBL:X82471; NID:g732858; PIDN:CA57854.1; PID:g732859
 A:Experimental source: strain HF159
 A:Accession: S61321
 A>Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 53-548 <LOM>
 A:Cross-references: EMBL:X82472; NID:g732852; PIDN:CA57855.1; PID:g732853
 A:Experimental source: strain SM1027
 C:Genetics:
 A:Gene: iga; NMA0905
 C:Superfamily: Iga-specific metalloendopeptidase
 C:Keywords: hydrolase; metalloproteinase

Query Match 100.0%; Score 558; DB 2; Length 1773;
 Best Local Similarity 100.0%; Pred. No. 1.6e-48;
 Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LYKKNRYALKSGSVNAPMPENGOTENNMDILMGSTOEAKKNAHKNRISGFSG 60
 DB 573 LYKKNRYALKSGSVNAPMPENGOTENNMDILMGSTOEAKKNAHKNRISGFSG 632
 QY 61 FFGENGKGHNGALNINFGKSAONRFLITGNTNGKISVTG 104
 DB 633 FFGENGKGHNGALNINFGKSAONRFLITGNTNGKISVTG 676

564

RESULT 3
 C81169
 IGA-specific metalloendopeptidase (EC 3.4.24.13) NMB0700 [imported] - *Neisseria meningitidis*
 C:Species: *Neisseria meningitidis*
 A:Variety: group B strain MDS8; strain 81139
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Dec-2000
 C:Accession: C81169; S61326
 R:Retelien, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Hart, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; et al.; Qiu, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizsa, M.
 Science 287, 1809-1815, 2000
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Vignani, A.
 A>Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
 A:Reference number: AB1000; MUID:20175755
 A:Accession: C81169
 A:Molecule type: DNA
 A:Residues: 1-1815 <GEN>
 A:Cross-references: GB:AE002424; GB:AE002098; NID:g7225923; PIDN:AAFA1117.1; PID:g722592
 A:Experimental source: serogroup B, strain MMD58
 R:Lombolt, H.; Poulsen, K.; Mogens, K.
 Mol. Microbiol. 15, 495-506, 1995
 A>Title: Comparative characterization of the iga gene encoding Iga1 protease in *Neisseria*
 A:Reference number: S61314; MUID:95302961
 A:Accession: S61326
 A>Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 53-548 <LOM>
 A:Cross-references: EMBL:X82477; NID:g732856; PIDN:CA57860.1; PID:g732857
 C:Genetics:
 A:Gene: NMB0700
 C:Superfamily: Iga-specific metalloendopeptidase
 C:Keywords: hydrolase; metalloproteinase

Query Match 99.3%; Score 554; DB 2; Length 1815;
 Best Local Similarity 99.0%; Pred. No. 4.1e-48;
 Matches 103; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LYKKNRYALKSGSVNAPMPENGOTENNMDILMGSTOEAKKNAHKNRISGFSG 60
 DB 573 LYKKNRYALKSGSVNAPMPENGOTENNMDILMGSTOEAKKNAHKNRISGFSG 632
 QY 61 FFGENGKGHNGALNINFGKSAONRFLITGNTNGKISVTG 104
 DB 633 FFGENGKGHNGALNINFGKSAONRFLITGNTNGKISVTG 676

RESULT 4
 A26039
 IGA-specific metalloendopeptidase (EC 3.4.24.13) precursor - *Neisseria gonorrhoeae* (s
 N:Alternate names: Iga protease; immunoglobulin A1 proteinase
 C:Species: *Neisseria gonorrhoeae*
 A:Variety: strain MS11
 C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 08-Dec-2000
 C:Accession: A26039; S09386
 R:Pohlner, J.; Halter, R.; Beyreuther, K.; Meyer, T.F.
 Nature 325, 458-462, 1987
 A>Title: Gene structure and extracellular secretion of *Neisseria gonorrhoeae* Iga prot.

A:Reference number: A26039; MUID:87115823
 A:Accession: A26039
 A:Molecule type: DNA
 A:Residues: 1-1532 <POH>
 A:Cross-references: GB:X04835; NID:g44868; PIDN:CA28558.1; PID:g44869
 A>Note: the authors translated the codon AAG for residue 668 as Asn
 R:Halter, R.; Pohlner, J.; Meyer, T.F.
 EMBO J. 8, 2737-2744, 1989
 A>Title: Mosaic-like organization of Iga protease genes in *Neisseria gonorrhoeae* gene
 A:Reference number: S09386; MUID:90060036
 A:Accession: S09386
 A>Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 281-325, 'N', 327-337, 'N', 339-427, 'W', 429-531, 'N', 533-615, 'V', 617-631, 'N', 6
 A:Experimental source: strain MS11
 C:Genetics:
 A:Gene: iga
 C:Superfamily: Iga-specific metalloendopeptidase
 C:Keywords: hydrolase; metalloproteinase; transmembrane protein
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:28-1533/Product: immunoglobulin A1 proteinase #status predicted <MAT>
 F:1986-987/Cleavage site: Pro-Ser (autolytic) #status predicted
 F:1018-1019/Cleavage site: Pro-Ser (autolytic) #status predicted
 F:1121-1122/Cleavage site: Pro-Ala (autolytic) #status predicted

Query Match 87.5%; Score 488; DB 2; Length 1532;
 Best Local Similarity 86.5%; Pred. No. 1.8e-41;
 Matches 90; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 LYKKNRYALKSGSVNAPMPENGOTENNMDILMGSTOEAKKNAHKNRISGFSG 60
 DB 584 LYKKNRYALKSGSVNAPMPENGOTENNMDILMGSTOEAKKNAHKNRISGFSG 643
 QY 61 FFGENGKGHNGALNINFGKSAONRFLITGNTNGKISVTG 104
 DB 644 FFGENGKGHNGALNINFGKSAONRFLITGNTNGKISVTG 687

RESULT 5
 A37023
 IGA-specific metalloendopeptidase (EC 3.4.24.13) type 1 precursor - *Haemophilus influ*
 N:Alternate names: immunoglobulin A1 proteinase type 1
 C:Species: *Haemophilus influenzae*
 C:Date: 31-Jan-1992 #sequence_revision 12-Jun-1992 #text_change 08-Dec-2000
 C:Accession: A37023
 R:Poulsen, K.; Brandt, J.; Hjorth, J.P.; Thøgersen, H.C.; Kilian, M.
 Infect. Immun. 57, 3097-3105, 1989

```
D6      :||||| :||| :||| ||||| ||||| :| :|  
        661 YFGEERK-NNGNLWTFKKSQNRFLLTGTLNGLDKVERG 703
```

TOXICITY OF THE INHIBITORS OF THE

RESULT 9

THIS PAGE IS BLANK (USPTO)

```

RESULT 2
ID 057035 PRELIMINARY; PRT; 992 AA.
AC 057035;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-NOV-1998 (TREMblrel. 08, last sequence update)
DE 01-OCT-2000 (TREMblrel. 15, last annotation update)
DE 1GAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=98010345; PubMed=9350862;
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J.-F.,
RA del Valle J., Achtmann M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread."
RL Mol. Microbiol. 25:1047-1064(1997).
RN 12
RP SEQUENCE OF 37-532 FROM N.A.
RC STRAIN=HF48;
RX MEDLINE=95302961; PubMed=7783620;
RA Lomholt H., Poulsen K., Mogens K.;
RT "Comparative characterization of the iga gene encoding IgA1 protease
RT in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
RT influenzae."
RL Mol. Microbiol. 15:495-506(1995).
DR EMBL: AF012211; AAC45794.1; -.
DR EMBL: X82480; CAA57863.1; -.
DR EMBL: X82475; CAA57858.1; -.
DR INTERPRO: IPR000710; -.
DR INTERPRO: IPR002195; -.
DR PRINTS: PRO0921; IGASERPTASE.
DR PROSITE: PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 992
SQ SEQUENCE 992 AA; 109228 MW; 3677DDE4CE6D9F69 CRC64;

Query Match 100.0%; Score 558; DB 2; Length 992;
Best Local Similarity 100.0%; Pred. No. 6.9e-51;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYKKNRYVALKSGSVNAPMPENGOTENNNDWILMGSTOEAKKNAMNHNORISGFSG 60
DB 557 LYKKNRYVALKSGSVNAPMPENGOTENNNDWILMGSTOEAKKNAMNHNORISGFSG 616
QY 61 FFEENGKGNGALNLFNGKSAQNRFLLTGCTNLNGKISVTG 104
DB 617 FFEENGKGNGALNLFNGKSAQNRFLLTGCTNLNGKISVTG 660

RESULT 3
Q9S6X5 PRELIMINARY; PRT; 992 AA.
ID 09S6X5;
AC 09S6X5;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DE 01-JUN-2000 (TREMblrel. 14, last annotation update)
DE 1GAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN 11
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J., del Valle J.,
RA Achtmann M.;

```

```

RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread."
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL: AF012203; AAC45786.1; -.
DR INTERPRO: IPR000710; -.
DR INTERPRO: IPR002195; -.
DR PRINTS: PRO0921; IGASERPTASE.
DR PROSITE: PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 992
SQ SEQUENCE 992 AA; 109239 MW; 306FF451B7213D49 CRC64;

Query Match 100.0%; Score 558; DB 2; Length 992;
Best Local Similarity 100.0%; Pred. No. 6.9e-51;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYKKNRYVALKSGSVNAPMPENGOTENNNDWILMGSTOEAKKNAMNHNORISGFSG 60
DB 557 LYKKNRYVALKSGSVNAPMPENGOTENNNDWILMGSTOEAKKNAMNHNORISGFSG 616
QY 61 FFEENGKGNGALNLFNGKSAQNRFLLTGCTNLNGKISVTG 104
DB 617 FFEENGKGNGALNLFNGKSAQNRFLLTGCTNLNGKISVTG 660

RESULT 4
Q9S6X4 PRELIMINARY; PRT; 992 AA.
ID 09S6X4;
AC 09S6X4;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)
DE 01-JUN-2000 (TREMblrel. 14, last annotation update)
DE 1GAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN 11
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J., del Valle J.,
RA Achtmann M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread."
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL: AF012204; AAC45787.1; -.
DR INTERPRO: IPR000710; -.
DR INTERPRO: IPR002195; -.
DR PRINTS: PRO0921; IGASERPTASE.
DR PROSITE: PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 992
SQ SEQUENCE 992 AA; 109239 MW; 9ED0A41019917CE3 CRC64;

Query Match 100.0%; Score 558; DB 2; Length 992;
Best Local Similarity 100.0%; Pred. No. 6.9e-51;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYKKNRYVALKSGSVNAPMPENGOTENNNDWILMGSTOEAKKNAMNHNORISGFSG 60
DB 557 LYKKNRYVALKSGSVNAPMPENGOTENNNDWILMGSTOEAKKNAMNHNORISGFSG 616
QY 61 FFEENGKGNGALNLFNGKSAQNRFLLTGCTNLNGKISVTG 104
DB 617 FFEENGKGNGALNLFNGKSAQNRFLLTGCTNLNGKISVTG 660

RESULT 5
O30575 PRELIMINARY; PRT; 997 AA.
ID O30575

```


AC 030575;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)
DE IGA1 PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98010345; PubMed=9350862;
RA Morelli G., Malorny B., Mueller K., Sella A., Wang J.-F.,
RA del Valle J., Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread";
RL MOL. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012210; AAC45793.1; -
DR INTERPRO; IPR000710; -
DR INTERPRO; IPR002195; -
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KM Protease.
FT NON_TER 1 1
FT SEQUENCE 997 AA; 110152 MW; 4DE40DE594E5E28E CRC64;

Query Match 100.0%; Score 558; DB 2; Length 997;
Best Local Similarity 100.0%; Pred. No. 6.9e-51;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYNNRYRYALKSGSVNAPMPENGOTENNNDWILMGSTOEBAKKNAMHKNRISGFSG 60
DB 558 LYNNRYRYALKSGSVNAPMPENGOTENNNDWILMGSTOEBAKKNAMHKNRISGFSG 617
DB 618 FFEENGKNGHNGALNPNKNGSAONRFLTGCTNLNGKISVTG 661

detected

RESULT 6
Q51169 PRELIMINARY; PRT; 1561 AA.
AC 051169;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)
DE IGA1 PROTEASE.
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HE13;
RX MEDLINE=95302961; PubMed=7783620;
RA Lombolt H., Poulsen K., Mogens K.;
RT "Comparative characterization of the iga gene encoding IGA1 protease
RT in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
RT influenzae";
RL MOL. Microbiol. 15:495-506(1995).
DR EMBL; X82474; CAA57857.1; -
DR MEROPS; S06.001; -
DR INTERPRO; IPR000710; -
DR INTERPRO; IPR002195; -
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KM Protease.
FT SEQUENCE 1561 AA; 171848 MW; 1C96E291A00017D5 CRC64;

Query Match 100.0%; Score 558; DB 2; Length 1561;

Best Local Similarity 100.0%; Pred. No. 1.2e-50;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYNNRYRYALKSGSVNAPMPENGOTENNNDWILMGSTOEBAKKNAMHKNRISGFSG 60
DB 584 LYNNRYRYALKSGSVNAPMPENGOTENNNDWILMGSTOEBAKKNAMHKNRISGFSG 643
QY 61 FFEENGKNGHNGALNPNKNGSAONRFLTGCTNLNGKISVTG 104
DB 644 FFEENGKNGHNGALNPNKNGSAONRFLTGCTNLNGKISVTG 687

RESULT 7
Q9JVB9 PRELIMINARY; PRT; 1773 AA.
AC 09JVB9;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)
DE IGA1 PROTEASE (EC 3.4.21.7).
GN IGA OR NMA0905.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=22491 / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrett B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491";
RL Nature 404:502-506(2000).
DR EMBL; AL162754; CAB84182.1; -
KM Protease; Hydrolase.
FT SEQUENCE 1773 AA; 196350 MW; CAC19E7313D76CE1 CRC64;

Query Match 100.0%; Score 558; DB 2; Length 1773;
Best Local Similarity 100.0%; Pred. No. 1.4e-50;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYNNRYRYALKSGSVNAPMPENGOTENNNDWILMGSTOEBAKKNAMHKNRISGFSG 60
DB 573 LYNNRYRYALKSGSVNAPMPENGOTENNNDWILMGSTOEBAKKNAMHKNRISGFSG 632
QY 61 FFEENGKNGHNGALNPNKNGSAONRFLTGCTNLNGKISVTG 104
DB 633 FFEENGKNGHNGALNPNKNGSAONRFLTGCTNLNGKISVTG 676

RESULT 8
Q9KOB4 PRELIMINARY; PRT; 1815 AA.
AC 09KOB4;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)
DE IGA-SPECIFIC SERINE ENDOPEPTIDASE.
GN NMB0700.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,

RA Nelson W.C., Gwin M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
 RA Cotton M.D., Ulterback T.R., Khouri H., Qin H., Vanathavan J.,
 RA Gill J., Scarlato V., Maignani V., Pizzo M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappunli R., Venter J.C.,
 RT "Complete genome sequence of *Neisseria meningitidis* serogroup B strain
 RT MC38.";
 RT Science 287:1809-1815(2000).
 DR EMBL: AE002424; AAF41117.1; -.
 DR TIGR: NMB0700; -.
 SQ SEQUENCE 1815 AA; 201077 MW; 2259D4D71762C57F CRC64;

Query Match 99.3%; Score 554; DB 2; Length 1815;
 Best Local Similarity 99.0%; Pred. No. 3.7e-50;
 Matches 103; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LYKKNRYIALKSGSVNAPMPENGOTENNMIWGSTOEAKKNAHKNORISGFSG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 573 LYKKNRYIALKSGSVNAPMPENGOTENNMIWGSTOEAKKNAHKNORISGFSG 632
 OY 61 FGEEGKGHNGALNLFNGKSAONRFLTTGGTNLNGKISVVG 104
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 633 FGEEGKGHNGALNLFNGKSAONRFLTTGGTNLNGKISVVG 676

RESULT 9
 O9S6X3 PRELIMINARY; PRT; 993 AA.

AC 09S6X3;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE IGAL PROTEASE PRECURSOR (FRAGMENT).
 GN IGA.
 OS *Neisseria meningitidis*.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; *Neisseria*.
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Morelli G., Malorny B., Mueller K., Sella A., Wang J., del Valle J.,
 RT "Clonal descent and microevolution of *Neisseria meningitidis* during 30
 RT years of epidemic spread.";
 RL Mol. Microbiol. 25:1047-1064(1997).
 DR EMBL: AF012208; AAC45791.1; -.
 DR INTERPRO: IPR000710; -.
 DR INTERPRO: IPR002195; -.
 DR PRINTS: PR00921; IGASERPTASE.
 DR PROSITE: PS00482; DIHYDROROTA_1; UNKNOWN_1.
 KW Protease.
 FT NON_TER 1
 FT NON_TER 993
 SQ SEQUENCE 993 AA; 109441 MW; 109FAA2EF88AC3C6 CRC64;

Query Match 96.8%; Score 540; DB 2; Length 993;
 Best Local Similarity 96.2%; Pred. No. 5.6e-49;
 Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 LYKKNRYIALKSGSVNAPMPENGOTENNMIWGSTOEAKKNAHKNORISGFSG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 558 LYKKNRYIALKSGSVNAPMPENGOTENNMIWGSTOEAKKNAHKNORISGFSG 617
 OY 61 FGEEGKGHNGALNLFNGKSAONRFLTTGGTNLNGKISVVG 104
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 618 FGEEGKGHNGALNLFNGKSAONRFLTTGGTNLNGKISVVG 661

RESULT 10
 O9S6X2 PRELIMINARY; PRT; 997 AA.

AC 09S6X2;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE IGAL PROTEASE PRECURSOR (FRAGMENT).
 GN IGA.
 OS *Neisseria meningitidis*.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; *Neisseria*.
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Morelli G., Malorny B., Mueller K., Sella A., Wang J., del Valle J.,
 RT "Clonal descent and microevolution of *Neisseria meningitidis* during 30
 RT years of epidemic spread.";
 RL Mol. Microbiol. 25:1047-1064(1997).
 DR EMBL: AF012209; AAC45792.1; -.
 DR INTERPRO: IPR000710; -.
 DR INTERPRO: IPR002195; -.
 DR PRINTS: PR00921; IGASERPTASE.
 DR PROSITE: PS00482; DIHYDROROTA_1; UNKNOWN_1.
 KW Protease.
 FT NON_TER 1
 FT NON_TER 997
 SQ SEQUENCE 997 AA; 109811 MW; 06F2E361E7E202E0 CRC64;

Query Match 96.8%; Score 540; DB 2; Length 997;
 Best Local Similarity 96.2%; Pred. No. 5.6e-49;
 Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 LYKKNRYIALKSGSVNAPMPENGOTENNMIWGSTOEAKKNAHKNORISGFSG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 558 LYKKNRYIALKSGSVNAPMPENGOTENNMIWGSTOEAKKNAHKNORISGFSG 617
 OY 61 FGEEGKGHNGALNLFNGKSAONRFLTTGGTNLNGKISVVG 104
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 DB 618 FGEEGKGHNGALNLFNGKSAONRFLTTGGTNLNGKISVVG 661

RESULT 11
 O57309 PRELIMINARY; PRT; 996 AA.

AC 057309;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE IGAL PROTEASE PRECURSOR (FRAGMENT).
 GN IGA.
 OS *Neisseria meningitidis*.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; *Neisseria*.
 OX NCBI_TaxID=487;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Morelli G., Malorny B., Mueller K., Sella A., Wang J.,
 RT "Clonal descent and microevolution of *Neisseria meningitidis* during 30
 RT years of epidemic spread.";
 RL Mol. Microbiol. 25:1047-1064(1997).
 RN [2]
 RP SEQUENCE OF 37-532 FROM N.A.
 RC STRAIN-ETH2;
 RX MEDLINE-95302961; PubMed-7783620;
 RA Lombolt H., Poulsen K., Mogens K.;
 RT "Comparative characterization of the iga gene encoding Iga1 protease
 RT in *Neisseria meningitidis*, *Neisseria gonorrhoeae* and *Haemophilus*
 RT influenzae.";
 RL Mol. Microbiol. 15:495-506(1995).
 DR EMBL: AF012207; AAC45790.1; -.
 DR EMBL: X82469; CA57852.1; -.
 DR EMBL: X82468; CA57851.1; -.
 DR INTERPRO: IPR000710; -.

DR INTERPRO: IPR002195; -
 DR PRINTS: PR00921; IGASERPTASE.
 DR PROSITE: PS004482; DIHYDROOROTASE_1; UNKNOWN_1.
 KM Protease.
 FT NON_TER 1
 FT NON_TER 996
 SQ SEQUENCE 996 AA; 109717 MW; AB39F7B98A41F986 CRC64;

Query Match 96.2%; Score 537; DB 2; Length 996;
 Best Local Similarity 95.2%; Pred. No. 1.2e-48;
 Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKNNRYVALKSGSYNAPPENGQETENDMDILMGSTOEAKKMANHKNRISGFSG 60
 DB 557 LYKNNRYVALKSGSYNAPPENGQETENDMDVPMGYKQEAQKMANHKNRISGFSG 616
 DB 617 FFEENGKGNHGNALNLFNPKSAQNRFLTGTLNKGKISVTG 104
 DB 617 FFEENGKGNHGNALNLFNPKSAQNRFLTGTLNKGKISVTG 660

RESULT 12
 ID 030574 PRELIMINARY; PRT; 996 AA.

AC 030574;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE ICA1 PROTEASE PRECURSOR (FRAGMENT).
 GN IGA.
 OS Neisseria meningitidis.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 NC NCBI_TaxID:487;
 RX MEDLINE-98010345; Pubmed-9350862;
 RA Morelli G., Malorny B., Mueller K., Sessler A., Wang J.-F.,
 RA del Valle J., Achman M.:
 RT "Clonal descent and microevolution of Neisseria meningitidis during 30
 RT years of epidemic spread."
 RL MOL. Microbiol. 25:1047-1064(1997).
 DR EMBL: AF012206; AAC45789.1;
 DR INTERPRO: IPR000710;
 DR INTERPRO: IPR002195;
 DR PRINTS: PR00921; IGASERPTASE.
 DR PROSITE: PS004482; DIHYDROOROTASE_1; UNKNOWN_1.
 KM Protease.
 FT NON_TER 1
 FT NON_TER 996
 SQ SEQUENCE 996 AA; 109688 MW; B3ABDEF54C337D9 CRC64;

Query Match 95.9%; Score 535; DB 2; Length 996;
 Best Local Similarity 95.2%; Pred. No. 1.9e-48;
 Matches 99; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LYKNNRYVALKSGSYNAPPENGQETENDMDILMGSTOEAKKMANHKNRISGFSG 60
 DB 557 LYKNNRYVALKSGSYNAPPENGQETENDMDVPMGYKQEAQKMANHKNRISGFSG 616
 QY 61 FFEENGKGNHGNALNLFNPKSAQNRFLTGTLNKGKISVTG 104
 DB 617 FFEENGKGNHGNALNLFNPKSAQNRFLTGTLNKGKISVTG 660

RESULT 13
 ID 09JXK3 PRELIMINARY; PRT; 1431 AA.
 AC 09JXK3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)

DE SERINE-TYPE PEPTIDASE.
 GN NMB1998.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 NC NCBI_TaxID:491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MC58 / SEROGROUP B;
 RX MEDLINE-20175755; Pubmed-10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., Deboy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
 RA Cotton M.D., Uppertack T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Maignani V., Piza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappelli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 RT MC58."
 RL Science 287:1809-1815(2000).
 DR EMBL: AE002549; AAF42325.1;
 DR TIGR: NMB1998;
 SQ SEQUENCE 1431 AA; 157645 MW; 6801F0E209C24EFC CRC64;

Query Match 22.9%; Score 128; DB 2; Length 1431;
 Best Local Similarity 25.5%; Pred. No. 4.1e-05;
 Matches 35; Conservative 22; Mismatches 44; Indels 36; Gaps 5;

QY 3 YKNRY-YYALKSGSYNAPPENGQETENDMDILMGSTOEAKKMANHKNRISGFS 59
 DB 555 HRNRRTYFLIKPGGNREFEPLN-MKNSTSWQITGNRRQAAQVQAENARDLITFG 613
 DB 614 GYLEENAGTGTAKAPSYKTEAELEKTRHIANAAVYGRPEYRGALNLHYRRTDSTL 673
 QY 60 GFEGE--ENGKG-----HNGALNLFNPKSAQNR 87
 DB 614 GYLEENAGTGTAKAPSYKTEAELEKTRHIANAAVYGRPEYRGALNLHYRRTDSTL 673
 QY 88 LITGTLNKGKISVTG 104
 DB 674 LITGTLNKGKISVTG 690

RESULT 14
 ID 09JWB4 PRELIMINARY; PRT; 1449 AA.

AC 09JWB4;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
 DE IGA-SPECIFIC SERINE ENDOPEPTIDASE (EC 3.4.21.72).
 GN IGA2 OR NMA0457.
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 NC NCBI_TaxID:65699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Z2491 / SEROGROUP A / SEROTYPE 4A;
 RX MEDLINE-20222556; Pubmed-10761919;
 RA Parkhill J., Achman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holroyd S.,
 RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria
 RT meningitidis Z2491."
 RL Nature 404:502-506(2000).
 DR EMBL: AL162753; CAB83754.1;
 KM Hydrolase.
 SQ SEQUENCE 1449 AA; 159070 MW; B858DB854BE50146 CRC64;

Query Match 20.6%; Score 115; DB 2; Length 1449;

Best Local Similarity 38.8%; Pred. No. 0.001;
Matches 26; Conservative 11; Mismatches 28; Indels 2; Gaps 1;

QY 38 TOEAKKNNAMHKNORISGFGFGEENGKHNALNENKSAQNRFLITGCTNLNG 97
DB 567 TOPSGKN--INRLNYSKEIAYNGWFGKDTTKTNGRLNLVYQPAEDRTLISGGTNLNG 624

QY 98 KISVYOG 104
DB 625 NITQTING 631

RESULT 15

09X7H1 PRELIMINARY; PRT; 1457 AA.
AC 09X7H1;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE APP PROTEIN.
GN APP.
OS Neisseria meningitidis serogroup B.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B:15:P1.16;
RA Abdel-Hadi H., Wooldridge K.G., Ala Aldeen D.A.;
RT "Identification and characterisation of neisserial App.";
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ242535; CAB43832.2; -;
DR INTERPRO: IPR000710; -;
DR PRINTS: PRO0921; IGASERPTASE.
DR PROSITE: PS00135; TRYPSIN_SER; UNKNOWN_1.
SQ SEQUENCE 1457 AA; 159956 MW; CE62390B6C04B781 CRC64;

Query Match 19.9%; Score 11; DB 2; Length 1457;

Best Local Similarity 29.6%; Pred. No. 0.0027;
Matches 24; Conservative 18; Mismatches 39; Indels 0; Gaps 0;

QY 24 NGQENNDMIIMGSTOEAKKNNAMHKNORISGFGFGEENGKHNALNENKSA 83
DB 558 NHNODKESTVITGKNKDIATGTGNNNSLDSKKEIAYNGWFGKDTTKTNGRLNLVYQPAE 617
QY 84 QNRFLITGCTNLNGKISVYOG 104
DB 618 DRTLISGGTNLNGNITQTING 638

Search completed: May 1, 2001, 15:04:51
Job time: 317 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 14:58:53 ; Search time 26.69 Seconds

(without alignments)
133.480 Million cell updates/sec

Title: US-09-142-970-1

Perfect score: 558

Sequence: 1 LYKKNRYALKSGSGSVNAP.....NRFLITGNTLNKISVTGG 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	488	87.5	1532	1	IGA_NEIGO
2	329.5	59.1	1541	1	IGA1_HAEIN
3	328.5	58.9	1694	1	IGA0_HAEIN
4	328.5	58.9	1702	1	IGA2_HAEIN
5	320.5	57.4	1545	1	IGA3_HAEIN
6	317.5	56.9	1849	1	IGA4_HAEIN
7	115.5	20.7	1409	1	HAP1_HAEIN
8	114	20.4	1394	1	HAP_HAEIN
9	94.5	16.9	989	1	PTP3_DICDI
10	82.5	14.8	451	1	ARP2_PLAFA
11	75.5	13.5	606	1	MANA_PIRSP
12	75.5	13.5	954	1	XYNA_RUMFL
13	74.5	13.4	719	1	ARP_YEAST
14	74.5	13.4	1596	1	NAM_DROME
15	72	12.9	712	1	TBPB_NEIMB
16	72	12.9	1556	1	GILTS_SYNY3
17	71.5	12.8	515	1	GUND_CLOCL
18	71.5	12.8	700	1	NONA_DROME
19	71	12.7	537	1	ARP_PLAFA
20	71	12.7	794	1	YB52_MYCPN
21	71	12.7	824	1	ROU_HUMAN
22	70.5	12.6	280	1	SMX1_SCHMA
23	70	12.5	591	1	Y040_MYCCE
24	69.5	12.5	641	1	LIP_STAHY
25	69	12.4	625	1	TBP2_HAEIN
26	69	12.4	1585	1	P3K3_DICDI
27	69	12.4	2334	1	MAPA_BACSU
28	68.5	12.3	101	1	VRP_SIVSP
29	68.5	12.3	447	1	SIF1_YEAST
30	68.5	12.3	569	1	MANC_PIRSP
31	68.5	12.3	571	1	MANB_PIRSP
32	68	12.2	464	1	DNA8_TREPA
33	68	12.2	563	1	SRCA_CHICK

34	68	12.2	1243	1	VG37_BPK3	O38394 bacterioph
35	67.5	12.1	1108	1	KB09_YEAST	P38080 saccharomyc
36	67.5	12.1	1584	1	RYK1_DICDI	P18160 dictyosteli
37	67	12.0	417	1	Y943_MERT1	O58355 methanococ
38	67	12.0	1080	1	HDC_DROME	O92408 drosophila
39	66.5	11.9	101	1	VRP_HV2D2	P15837 human immu
40	66.5	11.9	245	1	YF25_HAEIN	P71391 haemophilus
41	66.5	11.9	338	1	G3P_PHARH	O13507 pflaffia rho
42	66.5	11.9	716	1	PEP_DROME	P41073 drosophila
43	66	11.8	550	1	SRCA_HUMAN	O14247 homo sapien
44	66	11.8	982	1	CAPB_ANASP	P28594 anabaena sp
45	66	11.8	1656	1	OMP8_RICJA	O06653 r outer mem

ALIGNMENTS

RESULT	1	IGA_NEIGO	STANDARD:	PRT: 1532 AA.
AC	P09790:			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-MAR-1989 (Rel. 10, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	IGA-SPECIFIC SERINE ENDOPEPTIDASE PRECURSOR (EC 3.4.21.72) (IGA PROTEASE).			
DE	IGA.			
OS	Neisseria gonorrhoeae.			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=485;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	STRAIN=MS11;			
RX	MEDLINE=87115823; PubMed=3027577;			
RA	Pohlner J., Halter R., Beyreuther K., Meyer T.F.;			
RT	"Gene structure and extracellular secretion of Neisseria gonorrhoeae Iga protease.";			
RL	Nature 325:458-462(1987).			
RP	ACTIVE SITE.			
RX	MEDLINE=90154052; PubMed=2105953;			
RA	Bachovchin W.W., Plant A.G., Flentke G.R., Lynch M., Kettner C.A.;			
RT	"Inhibition of Iga1 proteinases from Neisseria gonorrhoeae and Hemophilus influenzae by peptide prolyl boronic acids.";			
RL	J. Biol. Chem. 265:3738-3743(1990).			
CC	- FUNCTION: THIS PROTEASE IS SPECIFIC FOR IMMUNOGLOBULIN A.			
CC	- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE SUBSTRATES ARE KNOWN.			
CC	- SUBCELLULAR LOCATION: SECRETED.			
CC	- DOMAIN: THE SIGNAL-PEPTIDE GUIDE THE PRECURSOR TO THE PERIPLASMIC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS.			
CC	- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASES).			
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib-sib.ch/announce/or_send_email_to_license@sib-sib.ch).			
CC	-----			
DR	EMBL; X04835; CAZ8538.1; -			
DR	PIR; A26039; A26039.			
DR	MEROPS; S06.001; -			
DR	InterPro: IPR000710; -			
DR	PRINTS; PR00921; IGASERPTASE.			
DR	Hydrolase; Serine protease; Zymogen; Autocatalytic cleavage;			
KW	Transmembrane; Signal.			
FT	SIGNAL 1 27			
FT	CHAIN 28 986 IGA-SPECIFIC SERINE ENDOPEPTIDASE.			

```

FT PROPEP 987 1532 HELPER PEPTIDE.
FT ACT_SITE 278 278 POTENTIAL.
FT SITE 987 987 CLEAVAGE (AUTO-).
FT SITE 1018 1019 CLEAVAGE (AUTO-).
FT SITE 1121 1122 CLEAVAGE (AUTO-).
SQ SEQUENCE 1532 AA; 168976 MW; 68FF4112BD22F40D CRC64;

Query Match
Best Local Similarity 87.5%; Score 488; DB 1; Length 1532;
Matches 90; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 1 LYKNTRYALKSGGSVNAPEPENGOTENNNDWILMGSTOEAKKNNKNNRISGFSG 60
DB 584 LYKNTRYALKSGGRNLNAPMENGVAENNNDWIFMGYTOEARKNNMNNRRIGDFG 643
QY 61 FPEENGKGHNGALNLFNGKSAQNRFLLTGTLNGKISVTG 104
DB 644 FPEENGKGHNGALNLFNGKSAQKRFLLTGALNKGKISVTG 687

RESULT 2
ID IGA1_HAEIN STANDARD; PRT; 1541 AA.
AC P42782;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (EC 3.4.21.72) (IGAL PROTEASE).
GN IGA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxId=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HK368 / SEROTYPE B;
RA MEDLINE=89379374; PubMed=2506130;
RA Poulsen K., Brandt J., Hjorth J.P., Thøgersen H.C., Killian M.;
RT "Cloning and sequencing of the immunoglobulin A1 protease gene (iga)
RT of Haemophilus influenzae serotype b.";
RL Infect. Immun. 57:3097-3105(1989).
RN [2]
RP MUTAGENESIS OF SER-288.
RC STRAIN-HK368 / SEROTYPE B;
RA MEDLINE=92234949; PubMed=1373717;
RA Poulsen K., Reinholdt J., Killian M.;
RT "A comparative genetic study of serologically distinct Haemophilus
RT influenzae type 1 immunoglobulin A1 proteases.";
RL J. Bacteriol. 174:2913-2921(1992).
CC -1- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
CC PRODUCING INTACT FC AND FAB FRAGMENTS.
CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT
CC CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE
CC SUBSTRATES ARE KNOWN.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X64357; CAA45708.1; -
CC DR EMBL: M87492; AAA24969.1; -

```

```

DR MEROPS; S06.001; -.
DR InterPro; IPR000710; -.
DR PRINTS; PRO00921; IGASERPTASE.
KW Hydrolase; Serine protease; Transmembrane; Zymogen; signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 1008 IMMUNOGLOBULIN A1 PROTEASE.
FT PROPEP 1009 1541 HELPER PEPTIDE (POTENTIAL).
FT ACT_SITE 288 288 PROBABLE.
FT MUTAGEN 288 288 S->T: LOSS OF ACTIVITY.
SQ SEQUENCE 1541 AA; 169370 MW; CE7257CB3196C600 CRC64;

Query Match
Best Local Similarity 59.1%; Score 329.5; DB 1; Length 1541;
Matches 59; Conservative 19; Mismatches 25; Indels 1; Gaps 1;

QY 1 LYKNTRYALKSGGSVNAPEPENGOTENNNDWILMGSTOEAKKNNKNNRISGFSG 60
DB 595 LNLNTYALRKGASTRSELPKNSGSENNMLYMGKTSDEAKRNNVHNNRMNGFNG 654
QY 61 FPEENGKGHNGALNLFNGKSAQNRFLLTGTLNGKISVTG 104
DB 655 YFGEDEK-NNGNLNTYFKGSEQNRFLLTGTLNGDLTVKRG 697

RESULT 3
ID IGA0_HAEIN STANDARD; PRT; 1694 AA.
AC P44969;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (EC 3.4.21.72) (IGAL PROTEASE).
GN IGA OR IGA1 OR HI0990.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxId=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SEROTYPE D;
RA WRIGHT A., Fishman Y., Tai F., Plant A.G.;
RT Submitted (May-1991) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KM20 / ATCC 51907;
RA MEDLINE=95350630; PubMed=7542800;
RA Felschmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Keriavage A.R., Butt C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uetshack T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
CC PRODUCING INTACT FC AND FAB FRAGMENTS.
CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT
CC CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE
CC SUBSTRATES ARE KNOWN.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

[illegible]

```
CC SIMILARITY) BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).
-----
CC CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL: M87489; AAA24966.1; -.
DR MEROPS; S06.001; -.
DR InterPro; IPR000710; -.
DR PRINTS; PR00921; IGASERPTASE.
KW Hydrolase; Serine protease; Transmembrane; zymogen; signal.
FT SIGNAL 1 25
FT CHAIN 26 1014 POTENTIAL.
FT PROPEP 1015 1702 IMMUNOGLOBULIN A1 PROTEASE.
FT ACT_SITE 288 288 HELPER PEPTIDE (POTENTIAL).
FT DOMAIN 1109 1124 PROBABLE.
FT FT 2 X 8 AA TANDEM REPEATS OF A-K-V-E-K-E-E-
FT REPEAT 1109 1116 K.
FT REPEAT 1117 1124 1.
FT REPEAT 1117 1124 2.
SQ SEQUENCE 1702 AA; 186539 MW; 860F70D2667807A6 CRC64;

Query Match 58.9%; Score 328.5; DB: 1; Length 1702;
Best Local Similarity 56.7%; Pred. NO. 7.5e-25;
Matches 59; Conservative 18; Mismatches 26; Indels 1; Gaps 1;

QY 1 LYYNRYRYALKSGSYNAVPENGGOTENNNDWIMLGSTOEAKRKNAMHKNNORISGFSG 60
   I :| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 601 LNLENTYTYALARKKASTSRSELPKSGESENEMWLMGKTSDDAKKNVNMHINNERNNGNG 660
   :| | | | | : | : | | | | | | | | | | | | | | | | | | | | | : |
   61 PFGEENGKGHGANGLNLFNPKNSAQNRLLNGTNMGKRISVTYG 104
   :| | | | | : | : | | | | | | | | | | | | | | | | | | | | | : |
Db 661 YFGEEGK-NNGNLNVTFEKKSQNRFLLTGTMNGDLKYEGK 703

RESULT 5
IGA3_HAEIN STANDARD; PRT; 1545 AA.
IGA3_HAEIN AC P45355;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (EC 3.4.21.72) (IGAI PROTEASE).
GN IGA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_taxid=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HK393 / NCTC 8467 / SEROTYPE B;
RX MEDLINE=92234949; Pubmed=1373717;
RA Poulsen K.; Reinholdt J.; Kilian M.;
RT "A comparative genetic study of serologically distinct Haemophilus
RT influenzae type 1 immunoglobulin A1 proteases." ;
RL J. Bacteriol. 174:2913-2921(1992).
CC -1- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
CC PRODUCING INTACT FC AND FAB FRAGMENTS.
CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT
CC CERTAIN PRO-I-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE
CC SUBSTRATES ARE KNOWN.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A POSE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROLYSIS (BY
CC SIMILARITY).
```


modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; U32710; -; NOT_ANNOTATED_CDS.

DR TIGR; H10248; -

DR InterPro: IPR000710; -

KM Hydrolyase; Serine protease; Transmembrane; Zymogen; Signal.

FT SIGNAL 1 25

FT CHAIN 26 ?

PROPEP ? 1409 ADHESION AND PENETRATION PROTEIN.

FT ACT_SITE 250 250 HELPER PEPTIDE (POTENTIAL).

FT ACCT_SITE 250 250 BY SIMILARITY.

SO SEQUENCE 1409 AA; 156797 MW; 63ABC93FA84D16E CRC64;

Query Match 20.7%; Score 115.5; DB 1; Length 1409;
Best Local Similarity 38.8%; Pred. No. 0.0009;
Matches 26; Conservative 12; Mismatches 28; Indels 1; Gaps 1;

OY 38 TOEAKKKAMNNKNNRISGSGFGEENGKNGALNLNNGKSAQNRFLITGNTING 97
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
DB 562 TAPSKKN-INKLDSKELAYNGWGETDKKNGRNLNLYKPTTEPTLLSGSTNLKG 620
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
OY 98 KISVTG 104
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
DB 621 DITQTKG 627

RESULT 8
ID HAP_HAEIN STANDARD; PRT: 1394 AA.
AC P45387;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE ADHESION AND PENETRATION PROTEIN PRECURSOR (EC 3.4.21.-).
GN HAF.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NHTI N187;
RA St Gene J.W. III, de la Morena M.L., Falkow S.;
RT "A Haemophilus influenzae Iga protease-like protein promotes intimate interaction with human epithelial cells."
RL Mol. Microbiol. 14:217-233(1994).
CC -1- FUNCTION: PROBABLE PROTEASE: PROMOTES ADHERENCE AND INVASION BY DIRECTLY BINDING TO A HOST CELL STRUCTURE.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: THE SIGNAL PEPTIDE GUIDE THE PRECURSOR TO THE PERIPLASMIC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U11024; AAB03707.1; -;
DR InterPro: IPR000710; -;
DR PRINTS; PR00921; IGASERPTASE.
KW Hydrolyase; Serine protease; Transmembrane; Zymogen; Signal.
FT SIGNAL 1 25

FT CHAIN 26 ? ADHESION AND PENETRATION PROTEIN.
FT PROPEP ? 1394 HELPER PEPTIDE (POTENTIAL).
FT ACT_SITE 243 243 BY SIMILARITY.
FT CONFLICT 1167 1167 MISSING (IN AAB03707).
SO SEQUENCE 1394 AA; 155441 MW; 5BF28660103F60F9 CRC64;

Query Match 20.4%; Score 114; DB 1; Length 1394;
Best Local Similarity 38.3%; Pred. No. 0.0013;
Matches 23; Conservative 13; Mismatches 24; Indels 0; Gaps 0;

OY 45 NAMNNKNNRISGSGFGEENGKNGALNLNNGKSAQNRFLITGNTINGKISVTG 104
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : : |
DB 551 NNINKLDYRKELAYNGWGETDKKNGRNLNLYKPTTEPTLLSGSTNLKGDTQTKG 610

RESULT 9
ID PTP3_DICDI STANDARD; PRT: 969 AA.
AC P54637;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROTEIN-TYROSINE PHOSPHATASE 3 (EC 3.1.3.48) (PROTEIN-TYROSINE-PHOSPHATE PHOSPHOHYDROLASE 3).
GN (PTPC1 OR PTP3) AND (PTPC2 OR PTP3).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AX3;
RX MEDLINE=96189126; PubMed=8628311;
RA Gampier M., Howard P.K., Hunter T., Firtel R.A.;
RT "Multiple roles of the novel protein tyrosine phosphatase PTP3 during Dictyostelium growth and development."
RL Mol. Cell. Biol. 16:2431-2444(1996).
CC -1- FUNCTION: SEEMS TO DEPHOSPHORYLATE A PROTEIN OF 130 KDA (P130).
CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: IN THE ANTERIOR-LIKE AND PRESTALK CELL TYPES.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT MODERATE LEVELS DURING GROWTH AND DEVELOPMENT.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-TYROSINE PHOSPHATASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; U38197; AAC47041.1; -;
DR HSP, P18031; IPTY.
DR Dictydb; DD01111; PTPC1.
DR Dictydb; DD07777; PTPC2.
DR InterPro: IPR000242; -;
DR InterPro: IPR000387; -;
DR Pfam: PF00102; Y_phosphatase_1.
DR PRINTS; PR00700; PTPPHPTASE.
DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW Hydrolyase.
FT ACT_SITE 649 649 BY SIMILARITY.
FT DOMAIN 64 71 POLY-ASN.
FT DOMAIN 109 118 POLY-ASN.
FT DOMAIN 137 190 POLY-ASN.
FT DOMAIN 249 257 POLY-SER.
FT DOMAIN 258 265 POLY-THR.

Db 209 HPATDTKKGQKFRREIOPSKSQDDRYSGFSGDDGGEYSNKNKSTLTDDQEGYGFSTNLEV 268
QY 77 NENGSAQNRFLLTGFTNLNGKISYTO 103
Db 269 DFHNKKLTGKLI RNNANTDNNQATTTO 295

Search completed: May 1, 2001, 15:03:34
Job time: 281 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 14:59:28 ; Search time 37.5 seconds
(without alignments)
53.278 Million cell updates/sec

Title: US-09-142-970-2

Perfect score: 562
Sequence: 1 LYKKNRYRYALKSGGSVNP.....NRFLLGTGTLNCKISVTGG 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfilest1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	505	89.9	1507	6 5268270-2	Patent No. 5268270
2	329.5	58.6	1541	5 PCT-US95-10661A-3	Sequence 3, Appl
3	322.5	57.4	1702	5 PCT-US95-10661A-5	Sequence 5, Appl
4	319.5	56.9	1545	5 PCT-US95-10661A-4	Sequence 4, Appl
5	312.5	55.6	1848	5 PCT-US95-10661A-6	Sequence 6, Appl
6	114	20.3	1394	5 PCT-US95-10661A-2	Sequence 2, Appl
7	83	14.6	1577	2 US-08-793-824-2	Sequence 2, Appl
8	76.5	13.6	631	1 US-08-487-890A-111	Sequence 11, App
9	76.5	13.6	631	2 US-08-478-435-111	Sequence 11, App
10	76.5	13.6	631	2 US-08-337-483-111	Sequence 11, App
11	76.5	13.6	631	2 US-08-478-373-111	Sequence 11, App
12	76.5	13.6	631	2 US-08-474-671-111	Sequence 11, App
13	76.5	13.6	631	3 US-08-483-577A-111	Sequence 11, App
14	74.5	13.3	430	1 US-08-945-848-8	Sequence 8, Appl
15	74.5	13.3	631	1 US-08-487-890A-115	Sequence 11, App
16	74.5	13.3	631	2 US-08-478-435-115	Sequence 11, App
17	74.5	13.3	631	2 US-08-337-483-115	Sequence 11, App
18	74.5	13.3	631	2 US-08-478-373-115	Sequence 11, App
19	74.5	13.3	631	3 US-08-474-671-115	Sequence 11, App
20	74.5	13.3	631	3 US-08-483-577A-115	Sequence 11, App
21	73.5	13.1	671	2 US-08-737-716-13	Sequence 13, Appl
22	71	12.6	708	3 US-08-613-009A-22	Sequence 22, Appl
23	71	12.6	806	1 US-07-980-528-2	Sequence 2, Appl
24	68	12.1	394	3 US-08-673-814-6	Sequence 6, Appl
25	68	12.1	648	1 US-08-487-890A-109	Sequence 109, App
26	68	12.1	648	2 US-08-478-435-109	Sequence 109, App
27	68	12.1	648	2 US-08-337-483-109	Sequence 109, App

28	68	12.1	648	2	US-08-478-373-109	Sequence 109, App
29	68	12.1	648	3	US-08-474-671-109	Sequence 109, App
30	68	12.1	648	1	US-08-483-577A-109	Sequence 109, App
31	67.5	12.0	613	3	US-08-272-875-3	Sequence 3, Appl
32	66.5	11.8	731	2	US-08-696-944-20	Sequence 20, Appl
33	66	11.7	434	2	US-08-795-475-3	Sequence 3, Appl
34	65	11.6	1529	2	US-08-728-470-10	Sequence 10, Appl
35	65	11.6	1600	2	US-08-617-697-10	Sequence 10, Appl
36	64.5	11.5	890	2	US-08-483-101-14	Sequence 14, Appl
37	64.5	11.5	1536	1	US-08-038-682-2	Sequence 2, Appl
38	64.5	11.5	1536	1	US-08-302-832-2	Sequence 2, Appl
39	64.5	11.5	1536	2	US-08-530-198-2	Sequence 2, Appl
40	64.5	11.5	1536	2	US-08-469-880-2	Sequence 2, Appl
41	64.5	11.5	1536	2	US-08-728-470-2	Sequence 2, Appl
42	64.5	11.5	1536	2	US-08-617-697-2	Sequence 2, Appl
43	64	11.4	711	3	US-08-613-009A-21	Sequence 21, Appl
44	64	11.4	1477	1	US-08-038-682-4	Sequence 4, Appl
45	64	11.4	1477	1	US-08-302-832-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
5268270-2
Patent No. 5268270
APPLICANT: Bayer, Thomas F., Halter, Roman, Pohlner, Johannes
TITLE OF INVENTION: PROCESS FOR PRODUCING PROTEINS USING GRAM
NEGATIVE HOST CELLS
NUMBER OF SEQUENCES: 6
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/171,872
FILING DATE: 01-JUL-1987
SEQ ID NO: 2:
LENGTH: 1507
5268270-2

Query Match 89.9%; Score 505; DB 6; Length 1507;
Best Local Similarity 88.5%; Pred. No. 7,7e-53;
Matches 92; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGGSVNPENGVTEINDMVFNGYQGEAKKNAMKKNORISGFSG 60
DB 559 LYKKNRYRYALKSGGRNALMPENGVAMENDMIFNGYQGEAKKNAMKKNRRIRIGDEGC 618

QY 61 FFEEENGKNGHGNALNLNENKSAQNRFLLTGTGTLNCKISVTGG 104
DB 619 FFEEENGKNGHGNALNLNENKSAQNRFLLTGTGTLNCKISVTGG 662

RESULT 2
PCT-US95-10661A-3
Sequence 3, Application PC/RUS9510661A
GENERAL INFORMATION:
APPLICANT: Washington University, et al.
TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10661A
FILING DATE: 16-AUG-1995


```
RESULT 5
PCT-US95-10661A-6
: Sequence 6, Application PC/TUS9510661A
: GENERAL INFORMATION:
: APPLICANT: Washington University, et al.
: TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
: STREET: 4 Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: United States
: ZIP: 94111-4187
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US95/10661A
: FILING DATE: 16-AUG-1995
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/296,791
: FILING DATE: 25-AUG-1994
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Treacartin, Richard F.
: REGISTRATION NUMBER: 31,801
: REFERENCE/DOCKET NUMBER: PP-59941/RFT
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 781-1989
: TELEFAX: (415) 398-3249
: TELEX: 910 277299
: INFORMATION FOR SEQ ID NO: 6:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1848 amino acids
: TYPE: amino acid
: TOPOLOGY: unknown
: PCT-US95-10661A-6

Query Match 55.6%; Score 312.5; DB 5; Length 1848;
Best Local Similarity 54.2%; Pred. No. 2.9e-29;
Matches 58; Conservative 16; Mismatches 28; Indels 3; Gaps 2;

QY 1 LYK-NRYALSGGSVNAPEMGVTENNNDVFMGYTOEAKKNAMHNQRTSGF 58
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 603 LYFNODNRSYTLTKKGASTRELPQNSGSENMNLYMGRTSDAKRNMVNNERNMGF 662
QY 59 SGFGEENGKG-HNGALNLNNGKSAONRFLLTGTNINNGKISYVG 104
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 663 NGYFGEERTKATQNGKLVNFGKSDONRFLLTGTNINGLDNLVEKG 709

RESULT 6
PCT-US95-10661A-2
: Sequence 2, Application PC/TUS9510661A
: GENERAL INFORMATION:
: APPLICANT: Washington University, et al.
: TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
: STREET: 4 Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: United States
: ZIP: 94111-4187
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10661A
FILING DATE: 16-AUG-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/296,791
FILING DATE: 25-AUG-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Treacartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: PP-59941/RFT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1394 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US95-10661A-2

Query Match 20.3%; Score 114; DB 5; Length 1394;
Best Local Similarity 38.3%; Pred. No. 3.1e-05;
Matches 23; Conservative 13; Mismatches 24; Indels 0; Gaps 0;

QY 45 NAMHKNQRTSGFSGFGEENGKGNALNLNNGKSAONRFLLTGTNINNGKISYVG 104
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 551 NNKLDYRKELAYNGWGETDKKHNGLNLTKPTEDRTLLSGTNLKGDITQKG 610

RESULT 7
US-08-793-824-2
: Sequence 2, Application US/08793824
: Patent No. 5981838
: GENERAL INFORMATION:
: APPLICANT: Simpson, Christine Lynn
: APPLICANT: Giffard, Philip Morrison
: APPLICANT: Jacques, Nicholas Anthony
: TITLE OF INVENTION: Genetic Manipulation of Plants to
: INCREASE STORED CARBOHYDRATES
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Griffith Hack & Co
: STREET: Level 8, 168 Walker Street
: CITY: New South Wales
: STATE: New South Wales
: COUNTRY: Australia
: ZIP: 2060
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/793,824
: FILING DATE:
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: AU PM7643
: FILING DATE: 24-AUG-1994
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 61 2 957 5944
: TELEFAX: 61 2 957 6288
: TELEX: 26547
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1577 amino acids
```

```

?      TYPE: amino acid
?      STRANDEDNESS:
?      TOPOLOGY: not relevant
?      MOLECULE TYPE: protein
?      ORIGINAL SOURCE:
?      ORGANISM: Streptococcus salivarius
US-08-793-824-2

```

Query Match	14.8%;	Score 83;	DB 2;	Length 1577;
Best Local Similarity	24.3%;	Pred. No. 0.22;		
Matches 28; Conservative	16;	Mismatches 41;	Indels 30;	Gaps 4;

```

QY      61  FEGEENGKGHNGALNLNFNCKSAQNFL-----LTGNTNLNGKI  99
      :  :  |  |  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db  1424  DYSQVKGAMANGRYDDGDSQAVSNQFIQIAANQWAYLNQDGHKYTGLOINNKV  1478

```

```

RESULT      8
US-08-487-890A-111
: Sequence 111, Application US/08487890A
: Patent No. 5708149
:
: GENERAL INFORMATION:
:   APPLICANT: Loosmore, Sheena
:   APPLICANT: Harkness, Robin
:   APPLICANT: Schryvers, Anthony
:   APPLICANT: Chong, Pele
:   APPLICANT: Gray-Owen, Scott
:   APPLICANT: Yang, Yan-Ping
:   APPLICANT: Murdin, Andrew
:   APPLICANT: Klein, Michel
:   TITLE OF INVENTION: Transferrin Receptor Genes
:   NUMBER OF SEQUENCES: 147
:
: CORRESPONDENCE ADDRESS:
:   ADDRESSEE: Slim & McBurney
:   STREET: 6th Floor, 330 University Avenue
:   CITY: Toronto
:   STATE: Ontario
:   COUNTRY: Canada
:   ZIP: M5G 1R7
:
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/08/487,890A
:   FILING DATE: 07-JUN-1993
:   CLASSIFICATION: 435
:
: PRIOR APPLICATION DATA:
:   APPLICATION NUMBER: US 08/175,116
:   FILING DATE: 29-DEC-1993
:
: PRIOR APPLICATION DATA:
:   APPLICATION NUMBER: US 08/148,968
:   FILING DATE: 08-NOV-1993
:   ATTORNEY/AGENT INFORMATION:
:     NAME: Stewart, Michael I
:     REGISTRATION NUMBER: 24,973
:     REFERENCE/DOCKET NUMBER: 1038-466 MIS:jb
:     TELECOMMUNICATION INFORMATION:
:       TELEPHONE: (416) 595-1155
:       TELEFAX: (416) 595-1163
:
: INFORMATION FOR SEQ ID NO: 111:
:   SEQUENCE CHARACTERISTICS:
:     LENGTH: 631 amino acids
:     TYPE: amino acid
:     TOPOLOGY: linear
:
: MOLECULE TYPE: protein
US-08-487-890A-111

```

Query Match	13.6%;	Score 76.5;	DB 1;	Length 631;
Best Local Similarity	25.8%;	Pred. No. 0.4;		
Matches 31; Conservative	7;	Mismatches 39;	Indels 43;	Gaps 5;

Qy 62 FGE-----ENGKGHGALNLFNGKSAQNRFLLTG-----GINLN 96
 || || || || ||
 Db 199 FGAYNRRSAISEDIDNLNNKNGAGLTSEFTVNFGTKR-----LTGKLYNERETNLN 252

```

RESULTS          9
US-08-478-435-111
Sequence 111, Application US/08478435
Patent No. 5922323
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena
APPLICANT: Harkness, Robin
APPLICANT: Schryvers, Anthony
APPLICANT: Chong, Pele
APPLICANT: Gray-Owen, Scott
APPLICANT: Yang, Yan-Ping
APPLICANT: Murgin, Andrew
APPLICANT: Klein, Michel
TITLE OF INVENTION: Transferrin Receptor Genes
NUMBER OF SEQUENCES: 147
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.23
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/478,435
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/337,483
FILING DATE: 08-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175,116
FILING DATE: 29-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,968
FILING DATE: 08-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-462 MIS:vg
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 111:
SEQUENCE CHARACTERISTICS:
LENGTH: 631 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-478-435-111

```

Query match . 13.6%; Score 76.5; DB 2; Length 631;


```

:
: GENERAL INFORMATION:
: APPLICANT: Loosmore, Sheena
: APPLICANT: Harkness, Robin
: APPLICANT: Schryvers, Anthony
: APPLICANT: Chong, Pele
: APPLICANT: Gray-Owen, Scott
: APPLICANT: Yang, Yan-ping
: APPLICANT: Murlin, Andrew
: APPLICANT: Klein, Michel
: TITLE OF INVENTION: Transferrin Receptor Genes
: NUMBER OF SEQUENCES: 147
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Slim & McBurney
: STREET: Suite 701, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/474,671
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/337,483
: FILING DATE: 08-NOV-1994
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/175,116
: FILING DATE: 29-DEC-1993
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/148,968
: FILING DATE: 08-NOV-1993
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Stewart, Michael I
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-465 MIS-VG
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 111:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 631 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-474-671-111
:
: Query Match 13.6%; Score 76.5; DB 3; Length 631;
: Best Local Similarity 25.8%; Pred. No. 0.4;
: Matches 31; Conservative 7; Mismatches 39; Indels 43; Gaps 5;
:
: QY 3 YKNRYVALKSGSVNAPMPENGVTENNDWVFMGYTOEAKKANMHNORISGSGF 61
: Db 150 YLGYGYAFYGNKTNLPVSGIAKYKGWDEITAT-----KNGQYSLFGSA 198
:
: QY 62 FGE-----ENKGHNGALNINFGKSAQNRFLLTG-----GTNIN 96
: Db 199 FGAYNRRSAISEDIDNLENNLKNAGLITSEFTVNFCTKK-----LTGKLYNERETNIN 252
:
: RESULT 13
: US-08-483-577A-111
: Sequence 111, Application US/08483577A
: Patent No. 6015688
: GENERAL INFORMATION:
: APPLICANT: Loosmore, Sheena
: APPLICANT: Harkness, Robin

```

```

:
: APPLICANT: Schryvers, Anthony
: APPLICANT: Chong, Pele
: APPLICANT: Gray-Owen, Scott
: APPLICANT: Yang, Yan-ping
: APPLICANT: Murlin, Andrew
: APPLICANT: Klein, Michel
: TITLE OF INVENTION: Transferrin Receptor Genes
: NUMBER OF SEQUENCES: 160
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Slim & McBurney
: STREET: Suite 701, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/483,577A
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/337,483
: FILING DATE: 08-NOV-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/175,116
: FILING DATE: 29-DEC-1993
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/148,968
: FILING DATE: 08-NOV-1993
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Stewart, Michael I
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-511
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 111:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 631 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-483-577A-111
:
: Query Match 13.6%; Score 76.5; DB 3; Length 631;
: Best Local Similarity 25.8%; Pred. No. 0.4;
: Matches 31; Conservative 7; Mismatches 39; Indels 43; Gaps 5;
:
: QY 3 YKNRYVALKSGSVNAPMPENGVTENNDWVFMGYTOEAKKANMHNORISGSGF 61
: Db 150 YLGYGYAFYGNKTNLPVSGIAKYKGWDEITAT-----KNGQYSLFGSA 198
:
: QY 62 FGE-----ENKGHNGALNINFGKSAQNRFLLTG-----GTNIN 96
: Db 199 FGAYNRRSAISEDIDNLENNLKNAGLITSEFTVNFCTKK-----LTGKLYNERETNIN 252
:
: RESULT 14
: US-08-945-848-8
: Sequence 84, Application US/08945848
: Patent No. 5968772
: GENERAL INFORMATION:
: APPLICANT: MATSUSHIRO, Aizo
: TITLE OF INVENTION: PEARL PROTEIN(NACREIN) AND PROCESS FOR
: TITLE OF INVENTION: THE SAME
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

```

Search completed: May 1, 2001, 14:59:29
Job time: 636 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 15:00:16 ; Search time 43.98 Seconds
(without alignments)
162.510 Million cell updates/sec

Title: US-09-142-970-2

Perfect score: 562

Sequence: 1 LYKKNRYRYALKSGSVNAP.....NRFLLTGNTLNKISVTG 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_67:*
1: pirl:*
2: pirl:*
3: pirl:*
4: pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	540	96.1	1561	2 S61314	IgA-specific metal
2	540	96.1	1773	2 A81937	IgA-specific metal
3	536	95.4	1815	2 C81169	IgA-specific metal
4	508	90.4	1532	2 A26039	IgA-specific metal
5	329.5	58.6	1541	2 A37023	IgA-specific metal
6	328.5	58.5	1694	2 H64106	IgA-specific metal
7	328.5	58.5	1702	2 A41859	IgA-specific metal
8	319.5	56.9	1545	2 B41859	IgA-specific metal
9	318.5	56.7	1849	2 C41859	IgA-specific metal
10	133	23.7	1431	2 A81018	serine-type peptid
11	115.5	20.6	709	2 C64057	IgA-specific metal
12	115	20.5	1449	2 B81963	IgA-specific serin
13	114	20.3	1394	2 D81019	IgA-specific serin
14	111.5	19.8	1457	2 D81019	adhesion and penet
15	83.5	14.9	451	2 A23535	clustered asparagi
16	83.5	14.9	465	2 S41644	glucosyltransferas
17	83	14.8	1577	2 T30858	polyadenylate-bind
18	82.5	14.7	243	2 T24981	hypothetical prote
19	77	13.7	163	2 T23076	hypothetical prote
20	77	13.7	625	2 D64107	transferrin-bindin
21	77	13.7	712	2 E61196	transferrin-bindin
22	77	13.7	982	1 A44831	phosphoenolpyruvat
23	77	13.7	1650	2 T18444	hypothetical prote
24	76	13.5	461	2 T10265	arabinogalactan-pr
25	76	13.5	631	2 S70908	transferrin-bindin
26	75.5	13.4	419	2 T40014	probable vesicular
27	75.5	13.4	484	2 G70846	hypothetical glyci
28	75.5	13.4	764	2 H71607	hypothetical prote
29	75	13.3	568	2 JC7210	molluscan shell ma

30	75	13.3	629	2 B75330	probable ribosomal
31	74.5	13.3	631	2 S70910	transferrin-bindin
32	74.5	13.3	719	2 S61046	ARPI protein - yea
33	74.5	13.3	1029	2 S64731	serine/threonine-s
34	74.5	13.3	1596	2 A33106	neurogenic locus m
35	74	13.2	607	2 T32968	hypothetical prote
36	74	13.2	1243	2 S07278	tail fiber protein
37	73.5	13.1	671	2 A38109	autolysin - Entero
38	72.5	12.9	642	2 D81401	probable flagellar
39	72.5	12.8	304	2 S65198	hypothetical prote
40	71.5	12.7	328	2 S24203	protein I - Neisse
41	71	12.6	402	2 T13614	N-acetyltransferas
42	71	12.6	697	2 T16306	hypothetical prote
43	71	12.6	806	2 S22765	heterogeneous ribo
44	71	12.6	2529	2 B64635	toxin-like outer m
45	70.5	12.5	125	2 T24982	hypothetical prote

ALIGNMENTS

RESULT 1
S61314
IgA-specific metalloendopeptidase (EC 3.4.24.13) homolog - Neisseria meningitidis
C:Species: Neisseria meningitidis
A:Variety: HF13
C:Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 08-Dec-2000
C:Accession: S61314
R:Holmolt, H.; Poulsen, K.; Mogens, K.
MOL: Microbiol. 15, 495-506, 1995
A:Title: Comparative characterization of the iga gene encoding IgA1 protease in Neiss
A:Reference number: S61314; MUID:95302961
A:Accession: S61314
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1561 <LOM>
A:Cross-references: EMBL:X82474; NID:9732873; PIDN:CA57857.1; PID:9732874
C:Superfamily: IgA-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase

Query Match 96.1%; Score 540; DB 2; Length 1561;
Best Local Similarity 96.2%; Pred. No. 9.2e-46;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
OY 1 LYKKNRYRYALKSGSVNAPMPENGVTENNDDVFMGYTQEEAKKNAMNNKQNRISGFSG 60
DB 584 LYKKNRYRYALKSGSVNAPMPENGOTENNDDVFMGYTQEEAKKNAMNNKQNRISGFSG 643
OY 61 FFEENGKNGHNCALNLNPNNGKSAQNRFLTLTGNTLNKISVTG 104
DB 644 FFEENGKNGHNCALNLNPNNGKSAQNRFLTLTGNTLNKISVTG 687
RESULT 2
A81937
IgA-specific metalloendopeptidase (EC 3.4.24.13) NMA0905 [imported] - Neisseria mening
N:Alternate names: IgA proteinase; IgA1 proteinase (EC 3.4.21.7) [misnomer]; immunogl
C:Species: Neisseria meningitidis
A:Variety: group A strain Z2491; strain HF117; strain HF159; strain SM1027
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 08-Dec-2000
C:Accession: A81937; S61317; S61318; S61321
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
R:Holmolt, H.; Jørgensen, K.; Leach, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
A:Reference number: A81775; MUID:20222556
A:Accession: A81937
A:Molecule type: DNA
A:Residues: 1-1773 <PAR>
A:Cross-references: GB:AL162754; GB:AL157959; NID:97379424; PIDN:CA884182.1; PID:9737
A:Experimental source: serogroup A, strain Z2491
R:Holmolt, H.; Poulsen, K.; Mogens, K.

Mol. Microbiol. 15, 495-506, 1995
A:Title: Comparative characterization of the iga gene encoding Iga1 protease in *Neisseria*
A:Reference number: S61314; MUID:95302961
A:Accession: S61317
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 53-548 <LOW>
A:Cross-references: EMBL:X82470; NID:g732854; PIDN:CAA57853.1; PID:g732855
A:Experimental source: strain HF117
A:Accession: S61318
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 53-548 <LOW>
A:Cross-references: EMBL:X82471; NID:g732858; PIDN:CAA57854.1; PID:g732859
A:Experimental source: strain HF159
A:Accession: S61321
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 53-548 <LOW>
A:Cross-references: EMBL:X82472; NID:g732852; PIDN:CAA57855.1; PID:g732853
A:Experimental source: strain SM1027
C:Genetics:
A:Gene: iga; NMA0905
C:Superfamily: Iga-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase

Query Match 96.1%; Score 540; DB 2; Length 1773;
Best Local Similarity 96.2%; Pred. No. 1.1e-45;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYYKNRYVALKSGSVNAPMPENGVTEENDVFMGYTOEAKKNNMHNKRNORISGFSG 60
|||||
DB 573 LYYKNRYVALKSGSVNAPMPENGVTEENDVFMGYTOEAKKNNMHNKRNORISGFSG 632
|||||

QY 61 FFEENGKGNHGNALNLFNGKSAQNRFLLTGTNLNGKISVTG 104
|||||
DB 633 FFEENGKGNHGNALNLFNGKSAQNRFLLTGTNLNGKISVTG 676
|||||

RESULT 3
C81169
Iga-specific metalloendopeptidase (EC 3.4.24.13) NMB0700 [Imported] - *Neisseria meningitidis*
C:Species: *Neisseria meningitidis*
A:Variety: group B strain MD58; strain 81139
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Dec-2000
C:Accession: C81169; S61326
R:Ettelein, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Halt, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignan, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grand, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappelli, R.; Ve
A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755
A:Accession: C81169
A:Molecule type: DNA
A:Residues: 1-1815 <TEU>
A:Cross-references: GB:AE002424; GB:AE002098; NID:g7225923; PIDN:AAF41117.1; PID:g722592
A:Experimental source: serogroup B, strain MMD58
R:Comblat, H.; Poulsen, K.; Mogens, K.
Mol. Microbiol. 15, 495-506, 1995
A:Title: Comparative characterization of the iga gene encoding Iga1 protease in *Neisseria*
A:Reference number: S61314; MUID:95302961
A:Accession: S61326
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 53-548 <LOW>
A:Cross-references: EMBL:X82477; NID:g732856; PIDN:CAA57860.1; PID:g732857
C:Genetics:
A:Gene: NMB0700
C:Superfamily: Iga-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase

Query Match 95.4%; Score 536; DB 2; Length 1815;
Best Local Similarity 95.2%; Pred. No. 2.7e-45;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYYKNRYVALKSGSVNAPMPENGVTEENDVFMGYTOEAKKNNMHNKRNORISGFSG 60
|||||
DB 573 LYYKNRYVALKSGSVNAPMPENGVTEENDVFMGYTOEAKKNNMHNKRNORISGFSG 632
|||||

QY 61 FFEENGKGNHGNALNLFNGKSAQNRFLLTGTNLNGKISVTG 104
|||||
DB 633 FFEENGKGNHGNALNLFNGKSAQNRFLLTGTNLNGKISVTG 676
|||||

RESULT 4
A26039
Iga-specific metalloendopeptidase (EC 3.4.24.13) precursor - *Neisseria gonorrhoeae* (s
N:Alternate names: Iga protease; immunoglobulin A1 proteinase
C:Species: *Neisseria gonorrhoeae*
A:Variety: strain MS11
C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 08-Dec-2000
C:Accession: A26039; S09386
R:Pohlner, J.; Halter, R.; Beyreuther, K.; Meyer, T.F.
Nature 325, 458-462, 1987
A:Title: Gene structure and extracellular secretion of *Neisseria gonorrhoeae* Iga prot
A:Reference number: A26039; MUID:87115823
A:Accession: A26039
A:Molecule type: DNA
A:Residues: 1-1532 <POH>
A:Cross-references: GB:X04835; NID:944868; PIDN:CAA28538.1; PID:944869
A:Note: The authors translated the codon AAG for residue 668 as Asn
R:Halter, R.; Pohlner, J.; Meyer, T.F.
EMBO J. 8, 2737-2744, 1989
A:Title: Mosaic-like organization of iga protease genes in *Neisseria gonorrhoeae* gene
A:Reference number: S09386; MUID:90060036
A:Accession: S09386
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 281-325, 'N', 327-337, 'N', 339-427, 'M', 429-531, 'N', 533-615, 'V', 617-631, 'N', 6
A:Experimental source: strain MS11
C:Genetics:
A:Gene: iga
C:Superfamily: Iga-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase; transmembrane protein
F:1-27/Domain: signal sequence #status predicted <Sig>
F:28-1532/Product: immunoglobulin A1 proteinase #status predicted <MAT>
F:986-987/Cleavage site: Pro-Ser (autolytic) #status predicted
F:1018-1019/Cleavage site: Pro-Ser (autolytic) #status predicted
F:1121-1122/Cleavage site: Pro-Ala (autolytic) #status predicted

Query Match 90.4%; Score 508; DB 2; Length 1532;
Best Local Similarity 88.5%; Pred. No. 1.4e-42;
Matches 92; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 LYYKNRYVALKSGSVNAPMPENGVTEENDVFMGYTOEAKKNNMHNKRNORISGFSG 60
|||||
DB 584 LYYKNRYVALKSGSVNAPMPENGVTEENDVFMGYTOEAKKNNMHNKRNORISGFSG 643
|||||

QY 61 FFEENGKGNHGNALNLFNGKSAQNRFLLTGTNLNGKISVTG 104
|||||
DB 644 FFEENGKGNHGNALNLFNGKSAQNRFLLTGTNLNGKISVTG 687
|||||

RESULT 5
A37023
Iga-specific metalloendopeptidase (EC 3.4.24.13) type 1 precursor - *Haemophilus influ*
N:Alternate names: immunoglobulin A1 proteinase type 1
C:Species: *Haemophilus influenzae*
C:Date: 31-Jan-1992 #sequence_revision 12-Jun-1992 #text_change 08-Dec-2000
C:Accession: A37023
R:Poulsen, K.; Brandt, J.; Hjorth, J.P.; Thogersen, H.C.; Killian, M.
Infect. Immun. 57, 3097-3105, 1989

Query Match	20.3%	Score 114	DB 2	Length 1394
Best Local Similarity	38.3%	Pred. No.	0.0024	
Matches	23	Conservative	24	Indels 0
		Mismatches		Gaps 0

RESULT 14

Query Match	19.8%	Score 11.5	DB 2	Length 1457
Best Local Similarity	34.2%	Pred. No. 0.0045		
Matches 27	Conservative 13	Mismatches 30	Indels 9	Gaps 2

```
QY      86  RELTGGTNLNGKISVTQG 104
          ||:||||||| | : ||
Db      620  TLLSGGTNLNGNITQTNG 638
```

RESULT	15
A23535	

Query Match	14.9%;	Score 83.5;	DB 2;	Length 451;
Best Local Similarity	25.0%;	Pred. No. 0.76;		
Matches 25; Conservative	9;	Mismatches 37;	Indels 29;	Gaps 4;

QY 65 ENKGHGHALNLTNGKSA-----QNRFLITGGTILN 96
| : : : : :
Db 204 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 243

Search completed: May 1, 2001, 15:00:17
Job time: 584 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 14:59:29 ; Search time 37.5 Seconds
(without alignments)
53.278 Million cell updates/sec

Title: US-09-142-970-3

Perfect score: 563
Sequence: 1 LYKKNRYALKSGSGSVNAP.....NRFLTGTLNKGKISVTGG 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCRNUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	492	87.4	1507	6 5268270-2	Patent No. 5268270
2	326.5	58.0	1545	5 PCT-US95-10661A-4	Sequence 4, Appli
3	320.5	56.9	1541	5 PCT-US95-10661A-3	Sequence 3, Appli
4	313.5	55.7	1702	5 PCT-US95-10661A-5	Sequence 5, Appli
5	303.5	53.9	1848	5 PCT-US95-10661A-6	Sequence 6, Appli
6	114	20.2	1394	5 PCT-US95-10661A-2	Sequence 2, Appli
7	78.5	13.9	430	2 US-08-945-848-8	Sequence 8, Appli
8	73.5	13.1	1577	2 US-08-793-824-2	Sequence 2, Appli
9	71	12.6	806	1 US-07-980-528-2	Sequence 6, Appli
10	70.5	12.5	631	1 US-08-487-890A-115	Sequence 115, App
11	70.5	12.5	631	2 US-08-478-435-115	Sequence 115, App
12	70.5	12.5	631	2 US-08-337-483-115	Sequence 115, App
13	70.5	12.5	631	2 US-08-474-671-115	Sequence 115, App
14	70.5	12.5	631	2 US-08-478-373-115	Sequence 115, App
15	70.5	12.5	631	2 US-08-474-671-115	Sequence 115, App
16	70.5	12.5	631	2 US-08-483-577A-115	Sequence 115, App
17	67.5	12.0	671	3 US-08-737-716-13	Sequence 13, App
18	66.5	11.8	890	2 US-08-483-101-14	Sequence 14, App
19	66.5	11.8	936	2 PCT-US93-05944-2	Sequence 2, Appli
20	65	11.5	648	1 US-08-487-890A-109	Sequence 109, App
21	65	11.5	648	2 US-08-478-435-109	Sequence 109, App
22	65	11.5	648	2 US-08-337-483-109	Sequence 109, App
23	65	11.5	648	2 US-08-478-373-109	Sequence 109, App
24	65	11.5	648	3 US-08-474-671-109	Sequence 109, App
25	65	11.5	648	3 US-08-483-577A-109	Sequence 109, App
26	65	11.5	1338	2 US-08-728-470-9	Sequence 9, Appli
27	65	11.5	1599	2 US-08-617-697-9	Sequence 9, Appli

28	64.5	11.5	631	1 US-08-487-890A-111	Sequence 111, App
29	64.5	11.5	631	2 US-08-478-435-111	Sequence 111, App
30	64.5	11.5	631	2 US-08-337-483-111	Sequence 111, App
31	64.5	11.5	631	2 US-08-478-373-111	Sequence 111, App
32	64.5	11.5	631	3 US-08-474-671-111	Sequence 111, App
33	64.5	11.5	631	3 US-08-483-577A-111	Sequence 111, App
34	64.5	11.5	691	1 US-08-064-174-2	Sequence 2, Appli
35	64.5	11.5	691	1 US-08-066-167-4	Sequence 4, Appli
36	64.5	11.5	691	2 US-08-449-733-2	Sequence 2, Appli
37	64.5	11.5	692	3 US-08-448-194-62	Sequence 62, Appli
38	64.5	11.5	711	1 US-08-487-890A-98	Sequence 98, Appli
39	64.5	11.5	711	2 US-08-478-435-98	Sequence 98, Appli
40	64.5	11.5	711	2 US-08-337-483-98	Sequence 98, Appli
41	64.5	11.5	711	2 US-08-478-373-98	Sequence 98, Appli
42	64.5	11.5	711	3 US-08-474-671-98	Sequence 98, Appli
43	64.5	11.5	711	3 US-08-483-577A-98	Sequence 98, Appli
44	64.5	11.5	711	3 US-08-448-194-8	Sequence 8, Appli
45	64	11.4	546	2 US-08-942-423-4	Sequence 4, Appli

ALIGNMENTS

```
RESULT 1
5268270-2
Patent No. 5268270
APPLICANT: Keyes, Thomas F.; Halter, Roman; Pohlner, Johannes
TITLE OF INVENTION: PROCESS FOR PRODUCING PROTEINS USING GRAM
NEGATIVE HOST CELLS
NUMBER OF SEQUENCES: 6
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/171,872
FILING DATE: 01-JUL-1987
SEQ ID NO: 2
LENGTH: 1507
5268270-2

Query Match 87.4% Score 492; DB 6; Length 1507;
Best Local Similarity 86.5% Pred. No. 2.9e-51;
Matches 90; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 LYKKNRYALKSGSVNAPPPENGQTFENNNDVPMGYSKOEAKNNKNNRISGSPG 60
|||
DB 559 LYKKNRYALKSGGRVAPPPENGVAENNNDWITMGITQEDAKNNKNNRIGDEGG 618
|||
QY 61 PFGEENGKNGALNLFNGKSAQNRFLTGTLNKGKISVTGG 104
|||
DB 619 PFDEENGKNGALNLFNGKSAQNRFLTGTLNKGKISVTGG 662

RESULT 2
PCT-US95-10661A-4
Sequence 4, Application PC/TUS9510661A
GENERAL INFORMATION:
APPLICANT: Washington University, et al.
TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Flehr, Hohbach, Teat, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PC/TUS95/10661A
FILING DATE: 16-AUG-1995
```



```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

Query Match      12.5%; Score 70.5; DB 1; Length 631;
Best Local Similarity 27.4%; Pred. No. 2;
Matches 29; Conservative 9; Mismatches 51; Indels 17; Gaps 5;

QY      2  YKKNRYIALSGGSVNAHPENGO-TENNNDVFMGYQOEAKKAMNHKN--ORISGF 58
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      153 FYSGGYGAAYFVGKOTATTLTGNGEATYKGTMSFTTAREGKNVSLFNNRGOAVSRSAT 212
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY      59 SGFGEENG-KGHGALNLNFGSKSAQRFLTG-----GTNIN 96
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      213 PGDIDLENGDAGLTSEFTVNFGTK-----LTGEPYNNRPTNIN 252
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

RESULT 12
US-08-478-435-115
; Sequence 115, Application US/08478435
; Patent No. 5922323
GENERAL INFORMATION:

```

RESULT 12
 US-08-478-435-115
 Sequence 115, Application US/08478435
 Patent No. 5922323
 GENERAL INFORMATION:
 APPLICANT: Loosmore, Sheena
 APPLICANT: Harkness, Robin
 APPLICANT: Schryvers, Anthony
 APPLICANT: Chong, Pele
 APPLICANT: Gray-Owen, Scott
 APPLICANT: Yang, Yan-Ping
 APPLICANT: Murdin, Andrew
 APPLICANT: Klein, Michel
 TITLE OF INVENTION: Transferrin Receptor Genes
 NUMBER OF SEQUENCES: 147
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sim & McBurney
 STREET: Suite 701, 330 University Avenue
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5G 1R7
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/478,435
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 15:00:17 ; Search time 43.98 Seconds
(Without alignments)
162.510 Million cell updates/sec

Title: US-09-142-970-3

Perfect score: 563
Sequence: 1 LYKNRYALKSGSVNAP.....NRFLLTGTLNGKISVTG 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	537	95.4	1561	2	S61314	Iga-specific metal
2	537	95.4	1773	2	A81937	Iga-specific metal
3	533	94.7	1815	2	C81169	Iga-specific metal
4	495	87.9	1532	2	A26039	Iga-specific metal
5	326.5	58.0	1545	2	B41859	Iga-specific metal
6	320.5	56.9	1541	2	A37023	Iga-specific metal
7	319.5	56.7	1694	2	H64106	Iga-specific metal
8	319.5	56.7	1702	2	A41859	Iga-specific metal
9	309.5	55.0	1849	2	C41859	Iga-specific metal
10	132	23.4	1431	2	A81018	serine-type peptid
11	114	20.2	1394	2	S60762	Iga-specific metal
12	113.5	20.2	1449	2	C64057	Iga-specific metal
13	112.5	20.0	1457	2	B81963	Iga-specific metal
14	109	19.4	1457	2	D81019	Iga-specific metal
15	86.5	15.4	451	2	A23535	clustered asparagi
16	80	14.2	1633	2	T23076	hypothetical prote
17	80	14.2	2529	2	B64635	toxlin-like outer m
18	79	14.0	179	2	T20206	hypothetical prote
19	79	14.0	402	2	T13614	N-acetyltransferas
20	79	14.0	629	2	B75330	probable ribosomal
21	79	14.0	1650	2	T18444	hypothetical prote
22	79	14.0	2399	2	H71879	toxlin-like outer m
23	78	13.9	461	2	T10265	arabinogalactan-pr
24	78	13.9	568	2	JC7210	molluscan shell ma
25	78	13.9	625	2	D64107	transferrin-bindin
26	78	13.9	1243	2	S07278	tail fiber protein
27	78	13.9	1577	2	T30858	glucosyltransferas
28	76.5	13.6	243	2	T24981	hypothetical prote
29	76.5	13.6	419	2	T40014	probable vesicular

Seq 3

30	76.5	13.6	954	1	S20907	endo-1,4-beta-xyla
31	76.5	13.6	1029	2	S64731	serine/threonine-s
32	75.5	13.4	465	2	S41644	polyadenylate-bind
33	75.5	13.4	484	2	G70846	hypothetical glyci
34	75	13.3	697	2	T16306	hypothetical prote
35	74.5	13.2	1596	2	A33106	neurogenic locus m
36	74	13.1	698	2	JH0163	No-on-transient A
37	74	13.1	700	2	JH0162	phosphoenolpyruvat
38	74	13.1	982	1	A44831	glucosyltransferas
39	74	13.1	1449	2	T30857	heat shock protein
40	73.5	13.1	350	2	T06689	heterogeneous ribo
41	73.5	13.1	806	2	S22765	asparagine-rich pr
42	73	13.0	537	2	A23770	glucosyltransferas
43	73	13.0	1449	2	T30552	C05B5.3 protein (c
44	72.5	12.9	364	2	S43574	nucleolin - chicke
45	72.5	12.9	694	1	DNCHNL	

ALIGNMENTS

RESULT 1
S61314
Iga-specific metalloendopeptidase (EC 3.4.24.13) homolog - Neisseria meningitidis
C:Species: Neisseria meningitidis
A:Variety: HF13
C>Date: 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 08-Dec-2000
R:Lomholt, H.; Poulsen, K.; Mogens, K.
Mol. Microbiol. 15, 495-506, 1995
A:Title: Comparative characterization of the iga gene encoding Iga1 protease in Neiss
A:Reference number: S61314; MUID:95302961
A:Accession: S61314
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1561 <Lom>
A:Cross-references: EMBL:X82474; NID:9732873; PIDN:CA57857.1; PID:9732874
C:Superfamily: Iga-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase

Query Match 95.4%; Score 537; DB 2; Length 1561;
Best Local Similarity 95.2%; Pred. No. 1.2e-45;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 LYKNRYALKSGSVNAPPENGOTENDWPMGYKQEEAQNAMNNKRNORISGFSG 60
|||||
Db 584 LYKNRYALKSGSVNAPPENGOTENDWIMLGSTQEEAKKNNKRNORISGFSG 643
|||||

OY 61 FFEENGKNGHNGALNLENKSAQNRFLTGTLNGKISVTG 104
|||||
Db 644 FFEENGKNGHNGALNLENKSAQNRFLTGTLNGKISVTG 687
|||||

RESULT 2
A81937
Iga-specific metalloendopeptidase (EC 3.4.24.13) NMA0905 [imported] - Neisseria menir
N:Alternate names: Iga1 proteinase; Iga1 proteinase (EC 3.4.21.7) [misnomer]; Immunog1
C:Species: Neisseria meningitidis
A:Variety: group A strain 22491; strain HF117; strain HF159; strain SM1027
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 08-Dec-2000
R:Parikh, J.; Achtmann, M.; James, K.D.; Bentley, S.D.; Churher, C.; Klee, S.R.; Mo
; Holroyd, S.; Jagels, K.; Leather, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491
A:Reference number: A81775; MUID:20222556
A:Accession: A81937
A:Molecule type: DNA
A:Residues: 1-1773 <PAR>
A:Cross-references: GB:AL162754; GB:AL157959; NID:97379424; PIDN:CAB84182.1; PID:97377
A:Experimental source: serogroup A, strain 22491
R:Lomholt, H.; Poulsen, K.; Mogens, K.

Mol. Microbiol. 15, 495-506, 1995
A>Title: Comparative characterization of the iga gene encoding Iga1 protease in *Neisseria*
A:Reference number: S61314; MUID:95302961
A:Accession: S61317
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 53-548 <LOW>
A:Cross-references: EMBL:X82470; NID:g732854; PIDN:CAA57853.1; PID:g732855
A:Experimental source: strain HF117
A:Accession: S61318
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 53-548 <LOW>
A:Cross-references: EMBL:X82471; NID:g732858; PIDN:CAA57854.1; PID:g732859
A:Experimental source: strain HF159
A:Accession: S61321
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 53-548 <LOW>
A:Cross-references: EMBL:X82472; NID:g732852; PIDN:CAA57855.1; PID:g732853
A:Experimental source: strain SM1027
C:Genetics:
A:Gene: iga
C:Superfamily: Iga-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase

Query Match 95.4%; Score 537; DB 2; Length 1773;
Best Local Similarity 95.2%; Pred. No. 1.4e-45;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYYKNRYALSKSGSVNAPMPENGOTENNDFVFGYKOEAKNMMKNNRISGSG 60
|||||
Db 573 LYYKNRYALSKSGSVNAPMPENGOTENNDFVFGYKOEAKNMMKNNRISGSG 632
|||||

QY 61 FGEEGKGHNGALNLFNGKSAQNRFLLTGTGTLNGKISYVG 104
|||||

Db 633 FGEEGKGHNGALNLFNGKSAQNRFLLTGTGTLNGKISYVG 676
|||||

RESULT 3
C81169
Iga-specific metalloendopeptidase (EC 3.4.24.13) NMB0700 [imported] - *Neisseria meningitidis*
C:Species: *Neisseria meningitidis*
A:Variety: group B strain MD58; strain 81139
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Dec-2000
C:Accession: C81169; S61326
R:Retelien, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Hart, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; R.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizsa, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V.
A>Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755
A:Accession: C81169
A:Molecule type: DNA
A:Residues: 1-1815 <TEP>
A:Cross-references: GB:AE002424; GB:AE002098; NID:g7225923; PIDN:AAFA1117.1; PID:g722592
A:Experimental source: serogroup B, strain MMD58
R:Tomholt, H.; Poulsen, K.; Mogens, K.
Mol. Microbiol. 15, 495-506, 1995
A>Title: Comparative characterization of the iga gene encoding Iga1 protease in *Neisseria*
A:Reference number: S61314; MUID:95302961
A:Accession: S61326
A>Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 53-548 <LOW>
A:Cross-references: EMBL:X82477; NID:g732856; PIDN:CAA57860.1; PID:g732857
C:Genetics:
A:Gene: NMB0700
C:Superfamily: Iga-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase

366

Query Match 94.7%; Score 533; DB 2; Length 1815;
Best Local Similarity 94.2%; Pred. No. 3.6e-45;
Matches 98; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYYKNRYALSKSGSVNAPMPENGOTENNDFVFGYKOEAKNMMKNNRISGSG 60
|||||
Db 573 LYYKNRYALSKSGSVNAPMPENGOTENNDFVFGYKOEAKNMMKNNRISGSG 632
|||||

QY 61 FGEEGKGHNGALNLFNGKSAQNRFLLTGTGTLNGKISYVG 104
|||||

Db 633 FGEEGKGHNGALNLFNGKSAQNRFLLTGTGTLNGKISYVG 676
|||||

RESULT 4
A26039
Iga-specific metalloendopeptidase (EC 3.4.24.13) precursor - *Neisseria gonorrhoeae* (s
N:Alternate names: Iga protease; immunoglobulin A1 proteinase
C:Species: *Neisseria gonorrhoeae*
A:Variety: strain MS11
C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 08-Dec-2000
C:Accession: A26039; S09386
R:Pohlner, J.; Halter, R.; Beyreuther, K.; Meyer, T.F.
Nature 325, 458-462, 1987
A>Title: Gene structure and extracellular secretion of *Neisseria gonorrhoeae* Iga prot
A:Reference number: A26039; MUID:87115823
A:Accession: A26039
A:Molecule type: DNA
A:Residues: 1-1532 <POH>
A:Cross-references: GB:X04835; NID:g44868; PIDN:CAA28538.1; PID:g44869
A:Note: The authors translated the codon AAG for residue 668 as Asn
R:Halter, R.; Pohlner, J.; Meyer, T.F.
EMBO J. 8, 2737-2744, 1989
A>Title: Mosaic-like organization of Iga protease genes in *Neisseria gonorrhoeae* gene
A:Reference number: S09386; MUID:90060036
A:Accession: S09386
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 281-325, 'N', 327-337, 'N', 339-427, 'N', 429-531, 'N', 533-615, 'V', 617-631, 'N', 6
A:Experimental source: strain MS11
C:Genetics:
A:Gene: iga
C:Superfamily: Iga-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase; transmembrane protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:128-1533/Product: immunoglobulin A1 proteinase #status predicted <MNT>
F:986-987/Cleavage site: Pro-Ser (autolytic) #status predicted
F:1018-1019/Cleavage site: Pro-Ser (autolytic) #status predicted
F:1121-1122/Cleavage site: Pro-Ala (autolytic) #status predicted

Query Match 87.9%; Score 495; DB 2; Length 1532;
Best Local Similarity 86.5%; Pred. No. 1.9e-41;
Matches 90; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 LYYKNRYALSKSGSVNAPMPENGOTENNDFVFGYKOEAKNMMKNNRISGSG 60
|||||
Db 584 LYYKNRYALSKSGSVNAPMPENGOTENNDFVFGYKOEAKNMMKNNRISGSG 643
|||||

QY 61 FGEEGKGHNGALNLFNGKSAQNRFLLTGTGTLNGKISYVG 104
|||||

Db 644 FGEEGKGHNGALNLFNGKSAQNRFLLTGTGTLNGKISYVG 687
|||||

RESULT 5
B41859
Iga-specific metalloendopeptidase (EC 3.4.24.13) type 1 precursor - *Haemophilus influ*
C:Species: *Haemophilus influenzae*
A:Variety: strain HK393
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 08-Dec-2000
C:Accession: B41859
R:Poulsen, K.; Reinholdt, J.; Killian, M.
J. Bacteriol. 174, 2913-2921, 1992

R:Wahlgren, M., Salslund, L.; Franzen, L.; Sundvall, M.; Wahlén, B.; Berins, K.; McNitt
Proc. Natl. Acad. Sci. U.S.A. 83, 2677-2681, 1986
A:Title: A plasmodium falciparum antigen containing clusters of asparagine residues.
A:Reference number: A23535; MUID:86206015
A:Accession: A23535
A:Molecule type: DNA
A:Residues: 1-451 <MAH>
A:Cross-references: GB: M13021; NID:g160079; PIDN:AAA29485.1; PID:g552179
R:Stolander, A.; Stahl, S.; Loygren, K.; Hansson, M.; Cavellier, L.; Waller, A.; Helmb
Exp. Parasitol. 76, 134-145, 1993
A:Title: Plasmodium falciparum: the immune response in rabbits to the clustered aspar
A:Reference number: A49245; MUID:93202225
A:Contents: 768
A:Accession: A49245
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 9-25 <SJO>
A:Cross-references: GB:S56699; NID:g298738; PID:g298739
A:Note: Sequence extracted from NCBI Backbone (NCBIN:127859, NCBI:P:127860)

	Query Match	15.4%	Score 86.5	DB 2	Length 451
	Best Local Similarity	28.4%	Pred. No. 0.35	Mismatches 9	Indels 33
	Matches 29	Conservative			Gaps
QY	5	NRYRYALKSGGSVAPMPENCG--TENNDWFMFYCKQEADKNMNMHKNNGRISGFSGF	62		
		: : : : : : : : : : : : : : : : : :			
Db	165	NYNFYNNNSSNNN-----NQQTNTQNN---FMNRNMK--NKNNNNNNNN-----	204		
QY	63	GEENGKGHGCAIINLFNGKSA-----ONRFLLTGSTLIN	96		
		: : : : : : : : : : : : : : : : :			
Db	205	--NNSNNNNMMANNFNENNNOOMNNPNPLNQNEFWLNNNNNNIN	243		

Search completed: May 1, 2001, 15:00:18
Job time: 585 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 15:03:35 ; Search time 26.69 seconds
(without alignments)
133.480 Million cell updates/sec

Title: US-09-142-970-3

Perfect score: 563

Sequence: 1 LYKRYRYALKSGSVNAP.....NRFLTGTNLNGKISVTQV 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	495	87.9	1532	IGA_NEIGO	P09790 neisseria g
2	326.5	58.0	1545	IGA3_HAEIN	P45385 haemophilus
3	320.5	56.9	1541	IGA1_HAEIN	P44782 haemophilus
4	319.5	56.7	1694	IGA0_HAEIN	P44969 haemophilus
5	319.5	56.7	1702	IGA2_HAEIN	P45384 haemophilus
6	309.5	55.0	1849	IGA4_HAEIN	P45386 haemophilus
7	114	20.2	1394	HAP_HAEIN	P45387 haemophilus
8	113.5	20.2	1409	HAP1_HAEIN	P44596 haemophilus
9	95.5	17.0	989	PTP3_DICDI	P54637 dictyosteli
10	86.5	15.4	451	TRP2_PLAFA	P13824 plasmodium
11	78	13.9	625	ABP2_HAEIN	P44971 haemophilus
12	76.5	13.6	1243	VG37_BPK3	C38394 bacterioph
13	76.5	13.6	1029	KSPI_YEAST	P29126 ruminooccu
14	74.5	13.2	1596	MAM_DROME	P38691 saccharomyc
15	74	13.1	700	NONA_DROME	P21519 drosophila
16	74	13.1	700	NONA_DROME	Q04047 drosophila
17	74	13.1	982	CAPP_ANASP	P28594 anabaena sp
18	73.5	13.1	824	ROU_HUMAN	Q00839 homo sapien
19	73	13.0	537	ARP_PLAFA	P04931 plasmodium
20	72.5	12.9	401	YKO3_CAEEL	P34291 caenorhabdi
21	72.5	12.9	694	NUCL_CHICK	P15771 gallus gall
22	72.5	12.9	719	ARP_YEAST	P32770 saccharomyc
23	72	12.8	563	SRCB_CHICK	Q01406 gallus gall
24	71	12.6	417	Y943_METJA	C58353 methanococc
25	70.5	12.5	507	VLL_HPV09	Q02480 human papil
26	70	12.4	550	SRCB_HUMAN	Q14247 homo sapien
27	70	12.4	712	TBPB_NEIMB	Q9K0V0 neisseria m
28	70	12.4	2334	WAPA_BACSU	Q07833 bacillus su
29	69.5	12.3	882	IF2_BOREU	O51741 borrelia bu
30	69	12.3	1113	N116_YEAST	Q02630 saccharomyc
31	68.5	12.2	606	MANA_PIRSP	P55296 piromyces s
32	67.5	12.0	671	ALYS_ENTFA	P37710 enterococcu
33	67.5	12.0	1902	P2P_LACPA	Q02470 lactobacill

ALIGNMENTS

RESULT 1	IGA_NEIGO	STANDARD;	PRT;	1532 AA.
ID	IGA_NEIGO			
AC	P09790;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-MAR-1989 (Rel. 10, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	IGA-SPECIFIC SERINE ENDOPEPTIDASE PRECURSOR (EC 3.4.21.72) (IGA PROTEASE).			
DE	PROTEASE).			
GN	IGA.			
OS	Neisseria gonorrhoeae.			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=485;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	STRAIN-MS11.			
RX	MEDLINE-87115823; PubMed-3027577;			
RA	Pollner J., Halter R., Beyreuther K., Meyer T.F.;			
RT	"Gene structure and extracellular secretion of Neisseria gonorrhoeae Iga protease."			
RL	Nature 325:458-462(1987).			
RN	[2]			
RP	ACTIVE SITE			
RX	MEDLINE-90154052; PubMed-2105953;			
RA	Bachovchin W.W., Plaut A.G., Flentke G.R., Lynch M., Kettner C.A.;			
RT	"Inhibition of Iga1 proteinases from Neisseria gonorrhoeae and Hemophilus influenzae by peptide prolyl boronic acids."			
RL	J. Biol. Chem. 265:3738-3743(1990).			
CC	-I- FUNCTION: THIS PROTEASE IS SPECIFIC FOR IMMUNOGLOBULIN A.			
CC	-I- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT CERTAIN PRO-I-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE SUBSTRATES ARE KNOWN.			
CC	-I- SUBCELLULAR LOCATION: SECRETED.			
CC	-I- DOMAIN: THE SIGNAL PEPTIDE GUIDE THE PRECURSOR TO THE PERIPLASMIC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE.			
CC	-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).			
CC	-I- This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; X04835; CAA28538.1; ..			
DR	PIR; A26039; A26039.			
DR	MEROPS; S06.001; ..			
DR	InterPro; IPR000710; ..			
DR	PRINTS; PR00921; IGASERPTASE.			
DR	Hydrolase; Serine protease; Zymogen; Autocatalytic cleavage; Transmembrane; Signal.			
KW	SIGNAL 1 27			
FT	CHAIN 28 986			
FT	IGA-SPECIFIC SERINE ENDOPEPTIDASE.			

```

Query Match          58.0%; Score 326.5; DB 1; Length 1545;
Best Local Similarity 57.5%; Pred. No. 1e-24;
Matches 61; Conservative 19; Mismatches 23; Indels 3; Gaps 2;

QY 1 LYY--KNRYRYALKSGSVNAPMPENGQTNNDWFMGYKOEAAQKNAWHKNNQRISGF 58
   ||: || ||||| ||: ||: || || || || || || || || || || || || ||
597 LYPNEENRTYALKDASIRSEFPQNRGESNNWLYMGTEKADAQKNAWHNNRNGF 656

QY 59 SGFPGENGKHGALNLFNGSKSAQNRFLLTGCTNLNGKISVTQG 104
   ||: ||||| || || || || || || || || || || || || || || || || ||
657 NGYFGEIEGR--NNGNLNVTFKSGEQNRFLLTGCTNLNGDLNVQQG 701

RESULT 3
IGAL_HAEIN
ID IGAL_HAEIN STANDARD; PRT; 1541 AA.
AC P42782;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (EC 3.4.21.72) (IGAL PROTEASE).
GN IGA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
[1]
RP SEQUENCE FROM N.A.
RP STRAIN=HK368 / SEROTYPE B;
RX MEDLINE=89379374; PubMed=2506130;
RA Poulsen K., Brandt J., Hjorth J.P., Thoegersen H.C., Kilian M.;
RT "Cloning and sequencing of the immunoglobulin A1 protease gene (iga)
of Haemophilus influenzae serotype b.;"
RL Infect. Immun. 57:3097-3105(1989).
[2]
RP MUTAGENESIS OF SER-288.
RP STRAIN=HK368 / SEROTYPE B;
RX MEDLINE=92234949; PubMed=1373717;
RA Poulsen K., Reinholdt J., Kilian M.;
RT "A comparative genetic study of serologically distinct Haemophilus
influenzae type 1 immunoglobulin A1 proteases.;"
RL J. Bacteriol. 174:2913-2921(1992).
CC -I- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
PRODUCING INTACT EC AND FAB FRAGMENTS.
CC -I- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT
CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE
SUBSTRATES ARE KNOWN.
CC -I- SUBCELLULAR LOCATION: SECRETED.
CC -I- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
SIMILARITY).
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY 'S6 (SERINE PROTEASE)".
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
CC EMBL; X64357; CAA45708.1; -
CC DDB; M87492; AAA24969.1; -
CC DR MEROPS; S06.001; -
CC InterPro; IPR000710; -
CC PRINTS; PR00921; IGASERPTASE.
KW Hydrolase; Serine protease; Transmembrane; Zymogen; Signal.
FT SIGNAL 1 25
FT CHAIN 26 1008 IMMUNOGLOBULIN A1 PROTEASE.

```

```
FT PROPEP 1009 1541 HELPER PEPTIDE (POTENTIAL).
FT ACT_SITE 288 288 PROBABLE.
FT MUTAGEN 288 288 S->T: LOSS OF ACTIVITY.
SQ SEQUENCE 1541 AA; 169370 MW; CE7257CB3196C600 CRC64;

Query Match 56.9%; Score 320.5; DB 1; Length 1541;
Best Local Similarity 54.8%; Pred. No. 4e-24;
Matches 57; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

QY 1 LYKKNRYALKSGGVNAPMPENGOTENNDFVFGYKOEBAQKAMNHKNQRTSGFSG 60
DB 595 LNLENTYIALRGASTRSELPKSGESNENMLYMGKTSDEAKRNMNHNERNMNGFNG 654
QY 61 FFEENGKNGHNGALNPNFGKSAQNRFLLTGTGTLNKGKTSVTOG 104
DB 655 YFGEERK-NNGNLNVTFKGEQNRFLLTGTGTLNGLDITVEKG 697

RESULT 4
ID IGA0_HAEIN STANDARD; PRT; 1694 AA.
AC P44969;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (EC 3.4.21.72) (IGAL PROTEASE).
GN IGA OR IGA1 OR HI0990.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN-SEROTYPE D;
RA Wright A., Fishman Y., Tai F., Plaut A.G.;
RL Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; Pubmed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
CC PRODUCING INTACT FC AND FAB FRAGMENTS.
CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT
CC CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE
CC SUBSTRATES ARE KNOWN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----NOT ANNOTATED_CDS.
DR EMBL: X59800; --
DR EMBL: U32779; AAC22651.1; --
DR MEROPS: S06.001; --
DR TIGR: HI0990; --
DR InterPro: IPR000710; --
DR PRINTS: PR00921; IGASERPTASE.
KW Hydrolase; Serine protease; Transmembrane; Zymogen; Signal.
FT SIGNAL 1 25
FT CHAIN 26 1014 IMMUNOGLOBULIN A1 PROTEASE.
FT PROPEP 1015 1694 HELPER PEPTIDE (POTENTIAL).
FT ACT_SITE 288 288 PROBABLE.
FT CONFLICT 253 254 EN -> GV (IN REF. 1).
FT CONFLICT 272 272 G -> A (IN REF. 1).
FT CONFLICT 464 464 G -> E (IN REF. 1).
FT CONFLICT 866 866 S -> T (IN REF. 1).
FT CONFLICT 1036 1036 A -> D (IN REF. 1).
FT CONFLICT 1074 1074 A -> G (IN REF. 1).
FT CONFLICT 1421 1421 H -> T (IN REF. 1).
FT CONFLICT 1545 1545 H -> T (IN REF. 1).
SQ SEQUENCE 1694 AA; 185539 MW; C52427013F93178C CRC64;

Query Match 56.7%; Score 319.5; DB 1; Length 1694;
Best Local Similarity 54.8%; Pred. No. 5.5e-24;
Matches 57; Conservative 20; Mismatches 26; Indels 1; Gaps 1;

QY 1 LYKKNRYALKSGGVNAPMPENGOTENNDFVFGYKOEBAQKAMNHKNQRTSGFSG 60
DB 601 LNLENTYIALRGASTRSELPKSGESNENMLYMGKTSDEAKRNMNHNERNMNGFNG 660
QY 61 FFEENGKNGHNGALNPNFGKSAQNRFLLTGTGTLNKGKTSVTOG 104
DB 661 YFGEERK-NNGNLNVTFKGEQNRFLLTGTGTLNGLDITVEKG 703

RESULT 5
ID IGA2_HAEIN STANDARD; PRT; 1702 AA.
AC P45384;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (EC 3.4.21.72) (IGAL PROTEASE).
GN IGA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN-HK715 / SEROTYPE B;
RX MEDLINE=92234949; Pubmed=1373717;
RA Poulisen K., Reinholdt J., Killian M.;
RT "A comparative genetic study of serologically distinct Haemophilus
RT influenzae type 1 immunoglobulin A proteases.";
RL J. Bacteriol. 174:2913-2921(1992).
CC -!- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
CC PRODUCING INTACT FC AND FAB FRAGMENTS.
CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT
CC CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE
CC SUBSTRATES ARE KNOWN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M87489; AAA24966.1; -
DR MEROPS; S06.001; -
DR InterPro; IPR000710; -
DR PRINTS; PR00921; IGASERPTASE.
KW Hydrolase; Serine protease; Transmembrane; Zymogen; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 1014 IMMUNOGLOBULIN A1 PROTEASE.
FT PROPEP 1015 1702 HELPER PEPTIDE (POTENTIAL).
FT ACT_SITE 288 288 PROBABLE.
FT DOMAIN 1109 1124 2 X 8 AA TANDEM REPEATS OF A-K-V-E-K-E-E-K.
FT REPEAT 1109 1116 1.
FT REPEAT 1117 1124 2.
SQ SEQUENCE 1702 AA; 186539 MW; 860F70D2667807A6 CRC64;

Query Match 56.78; Score 319.5; DB 1; Length 1702;
Best Local Similarity 54.88; Pred. No. 5.6e-24;
Matches 57; Conservative 20; Mismatches 26; Indels 1; Gaps 1;
QY 1 LYKNRYALKSGGSVNAAMPENGQTEENNDFVPMGYKQEAQKAMNHNKRNORISGFSG 60
DB 601 LNLENYIYALKRGASTRESELPKSGESNENWLYMGKTSDEAKRVNHNHINRMNGFNG 660
QY 61 FFGENGKGHGNALNPNKSAQNRELLTGGTNLNGKISVTQ 104
DB 661 YFGEPEGK-NNGNLNVTFKGSEQNRELLTGGTNLNGDLKVEKG 703

RESULT 6
ID IGAA_HAEIN STANDARD; PRT; 1849 AA.
AC P45386;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (EC 3.4.21.72) (IGAI PROTEASE).
GN IGA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NHTI HK61;
RX MEDLINE=92234949; PubMed=1373717;
RA Poulsen K., Reinholdt J., Kilian M.;
RT "A comparative genetic study of serologically distinct Haemophilus
RT influenzae type 1 immunoglobulin A1 proteases.";
RL J. Bacteriol. 174:2913-2921(1992).
CC -!- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
CC PRODUCING INTACT EC AND FAB FRAGMENTS.
CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT
CC CERTAIN PRO-|-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE
CC SUBSTRATES ARE KNOWN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE.
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M87491; AAA24968.1; -
DR MEROPS; S06.001; -
DR InterPro; IPR000710; -
DR PRINTS; PR00921; IGASERPTASE.
KW Hydrolase; Serine protease; Transmembrane; Zymogen; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 1021 IMMUNOGLOBULIN A1 PROTEASE.
FT PROPEP 1022 1849 HELPER PEPTIDE (POTENTIAL).
FT ACT_SITE 299 299 PROBABLE.
SQ SEQUENCE 1849 AA; 202957 MW; 79A7D018C7150AEA CRC64;

Query Match 55.08; Score 309.5; DB 1; Length 1849;
Best Local Similarity 53.38; Pred. No. 6e-23;
Matches 57; Conservative 19; Mismatches 28; Indels 3; Gaps 2;
QY 1 LYTK--NRYIYALKSGGSVNAAMPENGQTEENNDFVPMGYKQEAQKAMNHNKRNORISGF 58
DB 604 LYFNQDNRSYITLKGASTRESELPQSGESNENWLYMGRTSDEAKRVNHNHINRMNGF 663
QY 59 SGFFGENGKG-HGNALNPNKSAQNRELLTGGTNLNGKISVTQ 104
DB 664 NGYFGEETATQNGKLVNTPFNKSGDONRELLTGGTNLNGDLNVEKG 710

RESULT 7
HAP_HAEIN
ID HAP_HAEIN STANDARD; PRT; 1394 AA.
AC P45387;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 35, Last annotation update)
DE ADHESION AND PENETRATION PROTEIN PRECURSOR (EC 3.4.21.-).
GN HAP.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NHTI N187;
RX MEDLINE=95131744; PubMed=7830568;
RA St Gene J.W. III, de la Morena M.L., Falkow S.;
RT "A Haemophilus influenzae IGA protease-like protein promotes intimate
RT interaction with human epithelial cells.";
RL Mol. Microbiol. 14:217-233(1994).
CC -!- FUNCTION: PROBABLE PROTEASE; PROMOTES ADHERENCE AND INVASION BY
CC DIRECTLY BINDING TO A HOST CELL STRUCTURE.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- DOMAIN: THE SIGNAL PEPTIDE GUIDE THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL: U11024; AAB03707.1; -
InterPro: IPR000710; -
PRINTS; PR00921; IGASERPTASE.
KW Hydrolase; Serine protease; Transmembrane; Zymogen; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 ? ADHESION AND PENETRATION PROTEIN.

```

SQ SEQUENCE      1409 AA;    156797 MW;   63ABC893FA84D16E CRC64;

Query Match          20.2%; Score 113.5; DB 1; Length 1409;
Best Local Similarity 38.7%; Pred. No. 0.0013;
Matches 24; Conservative 11; Mismatches 22; Indels 5; Gaps 1;
```

```

RESULT 9
PTP3_DICDI
ID PTP3_DICDI STANDARD; PRT; 989 AA.
AC P34637;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROTEIN-TYROSINE PHOSPHATASE 3 (EC 3.1.3.48) (PROTEIN-TYROSINE-
DE PHOSPHATE PHOSPHOHYDROLASE 3).
DE (PTPCL OR PTP3) AND (PTPC2 OR PTP3).
GN Dictyostelium discoideum (Slime mold).
OS Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
NCBI_TaxID=44689;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=AX3;
RC MEDLINE=96189126; PubMed=8628311;
RX Gamber M., Howard P.K., Hunter T., Firtel R.A.;
RT "Multiple roles of the novel protein tyrosine phosphatase PTP3 during
RT Dictyostelium growth and development.";
RL Mol. Cell. Biol. 16:2431-2444(1996).
CC -1- FUNCTION: SEEMS TO DEPHOSPHORYLATE A PROTEIN OF 130 KDA (P130).
CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: IN THE ANTERIOR-LIKE AND PRESTALK CELL TYPES.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT MODERATE LEVELS DURING GROWTH
CC AND DEVELOPMENT.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC TYROSINE PHOSPHATASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).

```

DR	EMBL; U38197; AAC47041.1; -.
DR	HSP; P18031; IPTV.
DR	DictyDb; DD01111; ptpC1.
DR	DictyDb; DD0???; ptpC2.
DR	InterPro; IPR000242; -.
DR	InterPro; IPR000387; -.
DR	Pfam; PF00102; Y-phosphatase; 1.
DR	PRINTS; PS00700; PRTPHPHTASE.
DR	PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR	PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR	PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
KW	Hydrolase.
FT	ACT_SITE 649 649
FT	DOMAIN 64 71
FT	DOMAIN 109 118
FT	DOMAIN 137 190
FT	DOMAIN 249 257
FT	DOMAIN 258 265

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 14:59:30 ; Search time 37.5 Seconds
(without alignments)
53.278 Million cell updates/sec

Title: US-09-142-970-4
Perfect score: 569
Sequence: 1 LYKKNRYRYALKSGRLNAP.....NRFLLTGGANLGGNGRVPK 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	520	91.4	1507	6	Patent No. 5268270
2	281.5	49.5	1541	5	PCT-US95-10661A-3
3	275.5	48.4	1702	5	PCT-US95-10661A-5
4	270.5	47.5	1545	5	PCT-US95-10661A-4
5	264.5	46.5	1848	5	PCT-US95-10661A-6
6	91	16.0	1394	5	PCT-US95-10661A-2
7	76	13.4	1577	2	US-08-793-824-2
8	74.5	13.1	631	1	US-08-487-890A-115
9	74.5	13.1	631	2	US-08-478-435-115
10	74.5	13.1	631	2	US-08-337-483-115
11	74.5	13.1	631	2	US-08-478-373-115
12	74.5	13.1	631	3	US-08-474-671-115
13	74.5	13.1	631	3	US-08-483-577A-115
14	71	12.5	430	2	US-08-945-848-8
15	68.5	12.0	605	1	US-08-462-884A-3
16	68.5	12.0	605	1	US-08-461-881B-3
17	68.5	12.0	605	2	US-09-123-960-3
18	67.5	11.9	631	1	US-08-487-890A-111
19	67.5	11.9	631	2	US-08-478-435-111
20	67.5	11.9	631	2	US-08-337-483-111
21	67.5	11.9	631	2	US-08-478-373-111
22	67.5	11.9	631	3	US-08-474-671-111
23	67.5	11.9	631	3	US-08-483-577A-111
24	66.5	11.7	759	2	US-08-450-351-2
25	66.5	11.7	759	2	US-08-450-351-4
26	66.5	11.7	1115	3	US-08-323-477-2
27	65	11.4	265	3	US-08-483-577A-158

28	65	11.4	310	3	US-08-483-577A-157	Sequence 157, App
29	65	11.4	365	3	US-08-483-577A-156	Sequence 156, App
30	65	11.4	404	3	US-08-483-577A-155	Sequence 155, App
31	65	11.4	411	3	US-08-483-577A-154	Sequence 154, App
32	65	11.4	417	3	US-08-483-577A-153	Sequence 153, App
33	65	11.4	430	3	US-08-483-577A-152	Sequence 152, App
34	65	11.4	463	3	US-08-483-577A-151	Sequence 151, App
35	65	11.4	529	3	US-08-483-577A-150	Sequence 150, App
36	65	11.4	547	3	US-08-483-577A-149	Sequence 149, App
37	65	11.4	647	3	US-08-483-577A-148	Sequence 148, App
38	65	11.4	660	1	US-08-487-890A-8	Sequence 8, Appli
39	65	11.4	660	1	US-08-487-890A-10	Sequence 10, Appl
40	65	11.4	660	2	US-08-478-435-8	Sequence 8, Appli
41	65	11.4	660	2	US-08-478-435-10	Sequence 10, Appl
42	65	11.4	660	2	US-08-337-483-8	Sequence 8, Appli
43	65	11.4	660	2	US-08-337-483-10	Sequence 10, Appl
44	65	11.4	660	2	US-08-478-373-8	Sequence 8, Appli
45	65	11.4	660	2	US-08-478-373-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
5268270-2
; Patent No. 5268270
; APPLICANT: Meyer, Thomas F.; Halter, Roman; Pohlner, Johannes
; TITLE OF INVENTION: PROCESS FOR PRODUCING PROTEINS USING GRAM
; NEGATIVE HOST CELLS
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/171,872
; FILING DATE: 01-JUL-1987
; SEQ ID NO: 2:
; LENGTH: 1507
5268270-2

Query Match 91.4%; Score 520; DB 6; Length 1507;
Best Local Similarity 88.4%; Pred. No. 2.5e-52;
Matches 99; Conservative 2; Mismatches 1; Indels 10; Gaps 2;
QY 1 LYKKNRYRYALKSGRLNAPENGVAENNDWFMGYTQEEARKNNNNRRIGDPGG 60
Db 559 LYKKNRYRYALKSGRLNAPENGVAENNDWFMGYTQEEARKNNNNRRIGDEGG 618
QY 61 FFDEENGKGHGALNLFNPKSAQNRELLTGGANLNG-----GN----GRP 102
Db 619 FFDEENGKGHGALNLFNPKSAQNRELLTGGANLNGKISVTQGNVLLSGRP 670

RESULT 2
PCT-US95-10661A-3
; Sequence 3, Application PC/TUS9510661A
; GENERAL INFORMATION:
; APPLICANT: Washington University, et al.
; TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10661A
; FILING DATE: 16-AUG-1995

```
;
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,791
; FILING DATE: 25-AUG-1994
;
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: FP-59941/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1541 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; PCT-US95-10661A-3
;
; Query Match 49.5%; Score 281.5; DB 5; Length 1541;
; Best Local Similarity 48.2%; Pred. No. 2.1e-24;
; Matches 54; Conservative 17; Mismatches 30; Indels 11; Gaps 2;
;
; QY 1 LYXKNRYVALKSGGRLNAPMPENGVAENNDWVFMGYTQEEARKNAMNNKNNRRIGDFGG 60
; DB 595 LNLNYYTIALKKGASTRSELPKSGESNENWLYMGKTSDEAKRNVNMHNNRNGNG 654
;
; QY 61 PFDEENGKNGALNLFNGKSAQNRFLITGGANLNGG-----NGRP 102
; DB 655 YFGEDEGK-NNGNLNVTFKGSEQNRFLITGGTNGNLGDLTVKGTFLLSGRP 705
;
; RESULT 3
; PCT-US95-10661A-5
; Sequence 5, Application PC/TUS9510661A
; GENERAL INFORMATION:
; APPLICANT: Washington University, et al.
; TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10661A
; FILING DATE: 16-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: FP-59941/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1702 amino acids
; TYPE: amino acid
```

```
;
; TOPOLOGY: unknown
; PCT-US95-10661A-5
;
; Query Match 48.4%; Score 275.5; DB 5; Length 1702;
; Best Local Similarity 47.3%; Pred. No. 1.2e-23;
; Matches 53; Conservative 17; Mismatches 31; Indels 11; Gaps 2;
;
; QY 1 LYXKNRYVALKSGGRLNAPMPENGVAENNDWVFMGYTQEEARKNAMNNKNNRRIGDFGG 60
; DB 601 LNLNYYTIALKKGASTRSELPKSGESNENWLYMGKTSDAKRVNMHNNRNGNG 660
;
; QY 61 PFDEENGKNGALNLFNGKSAQNRFLITGGANLNGG-----NGRP 102
; DB 661 YFGEDEGK-NNGNLNVTFKGSEQNRFLITGGTNGNLGDLKVEKGTFLLSGRP 711
;
; RESULT 4
; PCT-US95-10661A-4
; Sequence 4, Application PC/TUS9510661A
; GENERAL INFORMATION:
; APPLICANT: Washington University, et al.
; TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10661A
; FILING DATE: 16-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: FP-59941/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1545 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; PCT-US95-10661A-4
;
; Query Match 47.5%; Score 270.5; DB 5; Length 1545;
; Best Local Similarity 48.2%; Pred. No. 4.1e-23;
; Matches 55; Conservative 16; Mismatches 30; Indels 13; Gaps 3;
;
; QY 1 LYXKNRYVALKSGGRLNAPMPENGVAENNDWVFMGYTQEEARKNAMNNKNNRRIGDF 58
; DB 597 LYFNEENRYTALKKXASIRSEFPQNGESNNNSWLYMGTEKADAKQKNNHNNRNGMGF 656
;
; QY 59 GGFDEENGKNGALNLFNGKSAQNRFLITGGANLNGG-----NGRP 102
; DB 657 NGYFGEDEGK-NNGNLNVTFKGSEQNRFLITGGTNGNLGDLNVQOQGTFLLSGRP 709
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1702 amino acids
; TYPE: amino acid
```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10661A
; FILING DATE: 16-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: FP-59941/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-10661A-2

Query Match 16.0%; Score 91; DB 5; Length 1394;
Best Local Similarity 32.4%; Pred. No. 0.036;
Matches 22; Conservative 10; Mismatches 26; Indels 10; Gaps

Qy 45 NAMNNKNNRRIGDGGFFDEENGKGHNGALNFNFGKSAQNRFLLTGGANLNGG----- 98
Db 551 NNINKLDYREIAYNGWFGTGDKNKNGRLNLIYKPTEDRTLLSSGTNLKGDITQTKG 610
Qy 99 ----NGRP 102
Db 611 KLFFSGRP 618

RESULT 7
US-08-793-824-2
; Sequence 2, Application US/08793824
; Patent No. 5981838
; GENERAL INFORMATION:
; APPLICANT: Simpson, Christine Lynn
; APPLICANT: Giffard, Philip Morrison
; APPLICANT: Jacques, Nicholas Anthony
; TITLE OF INVENTION: Genetic Manipulation of Plants to
; TITLE OF INVENTION: Increase Stored Carbohydrates
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Grifith Hack & Co
; STREET: Level 8, 168 Walker Street
; CITY: No. 5981838th Sydney
; STATE: New South Wales
; COUNTRY: Australia
; ZIP: 2060
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,824
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PM7643
; FILING DATE: 24-AUG-1994
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 61 2 9957 5944
; TELEFAX: 61 2 957 6288

```

```

; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 631 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-487-890A-115

Query Match 13.1%; Score 74.5; DB 1; Length 631;
Best Local Similarity 28.1%; Pred.No.1.1; Mismatches 8; Indels 17; Gaps 5;
Matches 32; Conservative

QY 2 YKNRYRYALKSGRLNADPENGVAE--NDWVFMGVTQEARKNAMNNKN--RRIGDF 58
DB 153 FYSGYGYAYYFGKQTATLTPVNGEATYKGTWSEFATATERGKKNYSLFNNRGQAYSRSAT 212
QY 59 GGFDFDEENG-KHNGALNLFNFKSAQNRFLLFG-----GANLNGGNGRPNK 104
DB 213 PGDIDLNGDAGLTSETVNFQKK-----LTGEPVYNERETNLNQSKDRKKH 260

RESULT 9
US-08-478-435-115
; Sequence 115, Application US/08478435
; Patent No. 5922323
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena
; APPLICANT: Harkness, Robin
; APPLICANT: Schryvers, Anthony
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Murdin, Andrew
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Transferrin Receptor Genes
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,435
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,483
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,116
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,968
; FILING DATE: 08-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-462 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 115:
; SEQUENCE CHARACTERISTICS:

```


LENGTH: 631 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-478-435-115

Query Match 13.1%; Score 74.5; DB 2; Length 631;
Best Local Similarity 28.1%; Pred. No. 1.1;
Matches 32; Conservative 8; Mismatches 57; Indels 17; Gaps 5;

QY 2 YKKNRYALKSGRLNAPENGVAE-NNDWVFMGYTQEEARKNMMNN--RRIGDF 58
Db 153 FSYGYGYAYFGKQTATTLFVNGEATYKGTWFSFITERGKNYSLFNNGQAYSRSAT 212
QY 59 GGFDEENG-KHGNGALNLFNGKSAQNRFLLTG-----LGEPPYNERETLNOSKDRKHK 104
Db 213 PGDIDLENGDAGLTSETVNFNGTK-----LGEPPYNERETLNOSKDRKHK 260

RESULT 10

US-08-337-483-115
Sequence 115, Application US/08337483
Patent No. 5922562

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena
APPLICANT: Harkness, Robin
APPLICANT: Schryvers, Anthony
APPLICANT: Chong, Pele
APPLICANT: Gray-Owen, Scott
APPLICANT: Yang, Yan-Ping
APPLICANT: Mordin, Andrew
APPLICANT: Klein, Michel

TITLE OF INVENTION: Transferrin Receptor Genes

NUMBER OF SEQUENCES: 147

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney

STREET: Suite 701, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/337,483

FILING DATE: 08-NOV-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-410 MIS:jb

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 115:

SEQUENCE CHARACTERISTICS:

LENGTH: 631 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-337-483-115

Query Match 13.1%; Score 74.5; DB 2; Length 631;
Best Local Similarity 28.1%; Pred. No. 1.1;
Matches 32; Conservative 8; Mismatches 57; Indels 17; Gaps 5;

QY 2 YKKNRYALKSGRLNAPENGVAE-NNDWVFMGYTQEEARKNMMNN--RRIGDF 58
Db 153 FSYGYGYAYFGKQTATTLFVNGEATYKGTWFSFITERGKNYSLFNNGQAYSRSAT 212
QY 59 GGFDEENG-KHGNGALNLFNGKSAQNRFLLTG-----LGEPPYNERETLNOSKDRKHK 104

Db 153 FSYGYGYAYFGKQTATTLFVNGEATYKGTWFSFITERGKNYSLFNNGQAYSRSAT 212
QY 59 GGFDEENG-KHGNGALNLFNGKSAQNRFLLTG-----GANLNGNGRPVK 104
Db 213 PGDIDLENGDAGLTSETVNFNGTK-----LGEPPYNERETLNOSKDRKHK 260

RESULT 11

US-08-478-373-115

Sequence 115, Application US/08478373

Patent No. 5922841

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena

APPLICANT: Harkness, Robin

APPLICANT: Schryvers, Anthony

APPLICANT: Chong, Pele

APPLICANT: Gray-Owen, Scott

APPLICANT: Yang, Yan-Ping

APPLICANT: Mordin, Andrew

APPLICANT: Klein, Michel

TITLE OF INVENTION: Transferrin Receptor Genes

NUMBER OF SEQUENCES: 147

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney

STREET: Suite 701, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/478,373

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/337,483

FILING DATE: 08-NOV-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/175,116

FILING DATE: 29-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/148,968

FILING DATE: 08-NOV-1993

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-463 MIS:vg

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 115:

SEQUENCE CHARACTERISTICS:

LENGTH: 631 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-478-373-115

Query Match 13.1%; Score 74.5; DB 2; Length 631;
Best Local Similarity 28.1%; Pred. No. 1.1;
Matches 32; Conservative 8; Mismatches 57; Indels 17; Gaps 5;

QY 2 YKKNRYALKSGRLNAPENGVAE-NNDWVFMGYTQEEARKNMMNN--RRIGDF 58
Db 153 FSYGYGYAYFGKQTATTLFVNGEATYKGTWFSFITERGKNYSLFNNGQAYSRSAT 212
QY 59 GGFDEENG-KHGNGALNLFNGKSAQNRFLLTG-----GANLNGNGRPVK 104

Db 213 PGDIDLENGDAGLTSEFTVNFCTKK-----LTGEPYNERETNLNQSKDRKHK 260

RESULT 12

US-08-474-671-115
; Sequence 115, Application US/08474671
; Patent No. 6008326

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena
APPLICANT: Harkness, Robin
APPLICANT: Schryvers, Anthony
APPLICANT: Chong, Pele
APPLICANT: Gray-Owen, Scott
APPLICANT: Yang, Yan-Ping
APPLICANT: Mordin, Andrew
APPLICANT: Klein, Michel

TITLE OF INVENTION: Transferrin Receptor Genes

NUMBER OF SEQUENCES: 147

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/474,671

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/337,483

FILING DATE: 08-NOV-1994

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/175,116

FILING DATE: 29-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/148,968

FILING DATE: 08-NOV-1993

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-465 MIS:vg

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 115:

SEQUENCE CHARACTERISTICS:

LENGTH: 631 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-474-671-115

Query Match 13.1%; Score 74.5; DB 3; Length 631;
Best Local Similarity 28.1%; Pred. No. 1.1;
Matches 32; Conservative 8; Mismatches 57; Indels 17; Gaps 5;

QY 2 YKKNRYVYALKSGGRNAPMPENGVAE--NNDWVFVGYTQEEARKNMNKN--RRIGDF 58

Db 153 FYSGYGYAYVFGKQTATTLPVNGEATYKGTWTSFATRGKNSYSLFNRRGOAYRRSAT 212

QY 59 GGFDEENG-KGHNGALNLFNGKSAQNRFLLTG-----CANLNGGNGRPVK 104

Db 213 PGDIDLENGDAGLTSEFTVNFCTKK-----LTGEPYNERETNLNQSKDRKHK 260

QY 59 GGFDEENG-KGHNGALNLFNGKSAQNRFLLTG-----CANLNGGNGRPVK 104

Db 213 PGDIDLENGDAGLTSEFTVNFCTKK-----LTGEPYNERETNLNQSKDRKHK 260

RESULT 13

US-08-483-577A-115
; Sequence 115, Application US/08483577A
; Patent No. 6015688

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena
APPLICANT: Harkness, Robin
APPLICANT: Schryvers, Anthony
APPLICANT: Chong, Pele
APPLICANT: Gray-Owen, Scott
APPLICANT: Yang, Yan-Ping
APPLICANT: Mordin, Andrew
APPLICANT: Klein, Michel

TITLE OF INVENTION: Transferrin Receptor Genes

NUMBER OF SEQUENCES: 160

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney

STREET: Suite 701, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/483,577A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/337,483

FILING DATE: 08-NOV-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/175,116

FILING DATE: 29-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/148,968

FILING DATE: 08-NOV-1993

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-511

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 115:

SEQUENCE CHARACTERISTICS:

LENGTH: 631 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-483-577A-115

Query Match 13.1%; Score 74.5; DB 3; Length 631;
Best Local Similarity 28.1%; Pred. No. 1.1;
Matches 32; Conservative 8; Mismatches 57; Indels 17; Gaps 5;

QY 2 YKKNRYVYALKSGGRNAPMPENGVAE--NNDWVFVGYTQEEARKNMNKN--RRIGDF 58

Db 153 FYSGYGYAYVFGKQTATTLPVNGEATYKGTWTSFATRGKNSYSLFNRRGOAYRRSAT 212

QY 59 GGFDEENG-KGHNGALNLFNGKSAQNRFLLTG-----CANLNGGNGRPVK 104

Db 213 PGDIDLENGDAGLTSEFTVNFCTKK-----LTGEPYNERETNLNQSKDRKHK 260

QY 59 GGFDEENG-KGHNGALNLFNGKSAQNRFLLTG-----CANLNGGNGRPVK 104

Db 213 PGDIDLENGDAGLTSEFTVNFCTKK-----LTGEPYNERETNLNQSKDRKHK 260

QY 59 GGFDEENG-KGHNGALNLFNGKSAQNRFLLTG-----CANLNGGNGRPVK 104

Db 213 PGDIDLENGDAGLTSEFTVNFCTKK-----LTGEPYNERETNLNQSKDRKHK 260

QY 59 GGFDEENG-KGHNGALNLFNGKSAQNRFLLTG-----CANLNGGNGRPVK 104

Db 213 PGDIDLENGDAGLTSEFTVNFCTKK-----LTGEPYNERETNLNQSKDRKHK 260

QY 59 GGFDEENG-KGHNGALNLFNGKSAQNRFLLTG-----CANLNGGNGRPVK 104

Db 213 PGDIDLENGDAGLTSEFTVNFCTKK-----LTGEPYNERETNLNQSKDRKHK 260

QY 59 GGFDEENG-KGHNGALNLFNGKSAQNRFLLTG-----CANLNGGNGRPVK 104

Db 213 PGDIDLENGDAGLTSEFTVNFCTKK-----LTGEPYNERETNLNQSKDRKHK 260

QY 59 GGFDEENG-KGHNGALNLFNGKSAQNRFLLTG-----CANLNGGNGRPVK 104

Db 213 PGDIDLENGDAGLTSEFTVNFCTKK-----LTGEPYNERETNLNQSKDRKHK 260

QY 59 GGFDEENG-KGHNGALNLFNGKSAQNRFLLTG-----CANLNGGNGRPVK 104

Db 213 PGDIDLENGDAGLTSEFTVNFCTKK-----LTGEPYNERETNLNQSKDRKHK 260

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 15:00:18 ; Search time 43.98 seconds
(without alignments)
162.510 Million cell updates/sec

Title: US-09-142-970-4
Perfect score: 569
Sequence: 1 LYKKNRYRYALKSGRLNAP.....NRFLTTGGANLNGNGRDPVK 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues
Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_67:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	523	91.9	1532	2 A26039	IgA-specific metal
2	458	80.5	1561	2 S61314	IgA-specific metal
3	458	80.5	1773	2 A81937	IgA-specific metal
4	454	79.8	1815	2 C81169	IgA-specific metal
5	281.5	49.5	1541	2 A37023	IgA-specific metal
6	281.5	49.5	1694	2 H64106	IgA-specific metal
7	281.5	49.5	1702	2 A41859	IgA-specific metal
8	270.5	47.5	1545	2 B41859	IgA-specific metal
9	270.5	47.5	1849	2 C41859	IgA-specific metal
10	134	23.6	1431	2 A81018	IgA-specific metal
11	91	16.0	1394	2 S60762	IgA-specific metal
12	90.5	15.9	1457	2 D81019	IgA-specific metal
13	89	15.6	709	2 C64057	IgA-specific metal
14	84.5	14.9	1449	2 B81963	IgA-specific metal
15	83	14.6	549	2 T20720	IgA-specific metal
16	80	14.1	461	2 T10265	IgA-specific metal
17	79.5	14.0	451	2 A23535	IgA-specific metal
18	79	13.9	1336	2 T18288	IgA-specific metal
19	78.5	13.8	419	2 T18420	IgA-specific metal
20	78.5	13.8	807	2 B71605	IgA-specific metal
21	78	13.7	568	2 JC7210	IgA-specific metal
22	78	13.7	1449	2 T30857	IgA-specific metal
23	76	13.4	163	2 T27368	IgA-specific metal
24	76	13.4	1577	2 T30858	IgA-specific metal
25	75.5	13.3	1538	2 H70846	IgA-specific metal
26	75.5	13.3	2150	2 S71629	IgA-specific metal
27	74.5	13.1	631	2 S70910	IgA-specific metal
28	74	13.0	290	2 T48274	IgA-specific metal
29	73.5	12.9	1596	2 A33106	IgA-specific metal

30	73	12.8	1381	2 E70806	hypothetical glyci
31	72.5	12.7	641	2 S57236	forked protein 2.5
32	72.5	12.7	1436	2 S57238	forked protein 5.4
33	72.5	12.7	1449	2 S57237	forked protein 5.6
34	72.5	12.7	1449	2 T30552	glucosyltransferas
35	72	12.7	132	2 T12108	glycine-rich prote
36	72	12.7	1029	2 S64731	serine/threonine-s
37	71.5	12.6	604	2 S39885	forked protein - f
38	71.5	12.6	1254	1 VHWVVE	structural polypro
39	71.5	12.6	1254	1 VHWVVT	structural polypro
40	71.5	12.6	1254	1 JQ1978	structural polypro
41	71.5	12.6	1255	1 B44213	structural polypro
42	71.5	12.6	1255	1 D44213	structural polypro
43	71.5	12.6	1538	2 S73296	glutamate synthase
44	71	12.5	210	2 F81386	probable acidic pe
45	71	12.5	232	2 T10646	hypothetical prote

ALIGNMENTS

RESULT 1
A26039
IgA-specific metalloendopeptidase (EC 3.4.24.13) precursor - Neisseria gonorrhoeae (s
N:Alternate names: IgA protease; immunoglobulin A1 proteinase
C:Species: Neisseria gonorrhoeae
A:Variety: strain MS11
C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 08-Dec-2000
C:Accession: A26039; S09386
R:Pohlner, J.; Haines, R.; Beyreuther, K.; Meyer, T.F.
Nature 325, 458-462, 1987
A:Title: Gene structure and extracellular secretion of Neisseria gonorrhoeae IgA prot
A:Reference number: A26039; MUID:87115823
A:Accession: A26039
A:Molecule type: DNA
A:Residues: 1-1532 <POH>
A:Cross-references: GB:X04835; NID:q44868; PIDN:CAA28538.1; PID:q44869
A:Note: The authors translated the codon AAG for residue 668 as Asn
R:Halter, R.; Pohlner, J.; Meyer, T.F.
EMBO J. 8, 2737-2744, 1989
A:Title: Mosaic-like organization of IgA protease genes in Neisseria gonorrhoeae gene
A:Reference number: S09386; MUID:90060036
A:Accession: S09386
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 281-325, N', 327-337, N', 339-427, 'M', 429-531, 'N', 533-615, 'V', 617-631, 'N', 6

C:Superfamily: IgA-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase; transmembrane protein
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-1532/Product: immunoglobulin A1 proteinase #status predicted <MAT>
F:986-987/Cleavage site: Pro-Ser (autolytic) #status predicted
F:1018-1019/Cleavage site: Pro-Ser (autolytic) #status predicted
F:1121-1122/Cleavage site: Pro-Ala (autolytic) #status predicted

Query Match 91.9%; Score 523; DB 2; Length 1532;
Best Local Similarity 88.4%; Pred. No. 1.9e-45;
Matches 99; Conservative 2; Mismatches 1; Indels 10; Gaps 2;

QY 1 LYKKNRYRYALKSGRLNAPMPNGVAENNDWTFMGYTQEEARKNNNNRRIGDFGG 60
|||||
Db 584 LYKKNRYRYALKSGRLNAPMPNGVAENNDWTFMGYTQEEARKNNNNRRIGDFGG 643
|||||

QY 61 FFDEENGKGHGALNLFNFGKSAQRNRLTGGANLNG-----GN-----GRP 102
|||||

Db 644 FFDEENGKGHGALNLFNFGKSAQRNRLTGGANLNGKISVTGNNVLLSGRP 695
|||||

RESULT 2
S61314

585

RESULT 3
A81937
Iga-specific metalloendopeptidase (EC 3.4.24.13) NMA0905 [imported] - Neisseria meningitidis
N:Alternate names: Iga proteinase; Iga1 proteinase (EC 3.4.21.7) [misnomer]; immunoglobulinase
C:Species: Neisseria meningitidis
A:Variety: group A strain Z2491; strain HF159; strain SM1027
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 08-Dec-2000
C:Accession: A81937; S61317; S61318; S61321
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556

A:Accession: S61318
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 53-548 <LOW>
A:Cross-references: EMBL:X82471; NID:g732858; PIDN:CAA57854.1; PID:g732859
A:Experimental source: strain HF159
A:Accession: S61321
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 53-548 <LOK>
A:Cross-references: EMBL:X82472; NID:g732852; PIDN:CAA57855.1; PID:g732853
A:Experimental source: strain SM1027
C:Genetics:
A:Gene: iqa; NMA0905

00002 JAN 00 23H0117Z /CCCT TBM CT 116

RESULT 4
C81169
IgA-specific metalloendopeptidase (EC 3.4.24.13)
C:Species: *Neisseria meningitidis*
A:Variety: group B strain MD58; strain 81139
C:Date: 31-Mar-2000 sequence revision 31-Mar-2000

A: Residues: 1-1815 <TEW>
A: Cross-references: GB:AE002424; GB:AE002098; NIH:100000000
A: Experimental source: serogroup B, strain MMD58
R: Lomholt, H.; Poulsen, K.; Mogensen, K.
Mol. Microbiol. 15, 495-506, 1995
A: Title: Comparative characterization of the Iga
A: Reference number: S61314; MUID:95302961
A: Accession: S61326
A: Status: preliminary; nucleic acid sequence not
A: Molecule type: DNA
A: Residues: 53-548 <LOW>
A: Cross-references: EMBL:X82477; NID:9732856; PIR:000000000
C: Genets:
A: Gene: NMB0700
C: Superfamily: Iga-specific metalloendopeptidase
C: Keywords: hydrolase; metalloproteinase

RESULT 5
A37023
IgA-specific metalloendopeptidase (EC 3.4.24.13)
N:Alternate names: immunoglobulin A1 proteinase
C:Species: Haemophilus influenzae
C:Date: 31-Jan-1992 #sequence_revision 12-Jun-1992
C:Accession: A37023
R:Poulsen, K.; Brandt, J.; Hjorth, J.P.; Thøgersen, I.
Infect. Immun. 57, 3097-3105, 1989

[illegible]

C41859
Iga-specific metalloendopeptidase (EC 3.4.24.13) type 1 precursor - Haemophilus influenzae
C:Species: Haemophilus influenzae
A:Variety: strain HK613
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 08-Dec-2000
C:Accession: C41859
R:Poulsen, K.; Reinholdt, J.; Killian, M.
J. Bacteriol. 174, 2913-2921, 1992
A:Title: A comparative genetic study of serologically distinct Haemophilus influenzae type 1
A:Reference number: A41859; MUID:92234949
A:Accession: C41859
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-1849 <POU>
A:Experimental source: strain HK613
A:Note: sequence extracted from NCBI backbone (NCBIP:97285)
C:Superfamily: Iga-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase

Query Match 47.5%; Score 270.5; DB 2; Length 1849;
Best Local Similarity 47.0%; Pred. No. 1.7e-19;
Matches 54; Conservative 15; Mismatches 33; Indels 13; Gaps 3;

Qy 1 LYK--NYRYALKSGRLNAPMPENGVAENNDVFMGYTOEARKNAMNNKNNRRIGDF 58
Db 604 LYFNQDNRSYTLKKGASTRSELPQNSGESNENWLYMGRTSDAKRNVNMHINERMNGF 663

Qy 59 GGFDEENGKG--HNGALNLFNGKSAQRNRLTGGANLNGG-----NGRP 102
Db 664 NGYFGEETKATQNGKLVNTFNGKSDQNRFLTGGTNGLDNLVKEGTLFLSGRP 718

RESULT 10
A01018
serine-type peptidase NMB1998 [imported] - Neisseria meningitidis (group B strain MD58)
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Dec-2000
C:Accession: A01018
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Qi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizzia, M. Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Venter, A.; Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Qi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizzia, M. Science 287, 1809-1815, 2000
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MD58.
A:Reference number: A81000; MUID:20175755
A:Accession: A81018
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1431 <TET>
A:Cross-references: GB:AE002549; GB:AE002098; NID:97227256; PIDN:AAF42325.1; PID:9722726
A:Experimental source: serogroup B, strain MMD58
C:Genetics:
A:Gene: NMB1998

Query Match 23.6%; Score 134; DB 2; Length 1431;
Best Local Similarity 28.8%; Pred. No. 1.3e-05;
Matches 42; Conservative 18; Mismatches 40; Indels 46; Gaps 7;

Qy 3 YKNYR--YALKSGRLNAPMPENGVAENNDVFMGYTOEARKNAMNNKNNR-RIGDFG 59
Db 555 HNRRTDYFILKPGNGPREFFPLN-MKNSTSQFIGNRQQAQEAQVARNAPDLITFG 613

Qy 60 GPFDE--ENGKG-----HNGALNLFNGKSAQRNRF 87
Db 614 GYLGENAOTGAAPSYSKTNEAIEKTHIANAAVYGRPEYRNGALNLHVRPKRTDSTL 673

Qy 88 LITGGANL-----CGN-----GRVP 103
Db 674 LLGGCMNLNGEVLIEGGNNIVSGREV 699

RESULT 11
S60762
Iga-specific serine endopeptidase (EC 3.4.21.72) precursor - Haemophilus influenzae (C:Species: Haemophilus influenzae
N:Alternate names: adhesion and penetration protein hap
C:Species: Haemophilus influenzae
C:Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 08-Dec-2000
C:Accession: S60762; S78001
R:St Geme III, J.W.; de la Morena, M.L.; Falkow, S. Mol. Microbiol. 14, 217-233, 1994
A:Title: A Haemophilus influenzae Iga protease-like protein promotes intimate interaction with host cells
A:Reference number: S60762; MUID:95131744
A:Accession: S60762
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-1394 <STG>
A:Cross-references: EMBL:U11024
A:Experimental source: strain N187
A:Accession: S78001
A:Molecule type: protein
A:Residues: 27-33 <ST2>
C:Genetics:
A:Gene: hap
C:Superfamily: Iga-specific metalloendopeptidase
C:Keywords: hydrolase; serine proteinase
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-1394/Product: Iga-specific serine endopeptidase #status experimental <MAT>
F:243/Active site: Ser #status predicted

Query Match 16.0%; Score 91; DB 2; Length 1394;
Best Local Similarity 32.4%; Pred. No. 0.31;
Matches 22; Conservative 10; Mismatches 26; Indels 10; Gaps 1;

Qy 45 NAMNNKNNRRICDGGFDEENGKNGALNLFNGKSAQRNRLTGGANLNGG-----98
Db 551 NNINKLDYRKAIYANGWFGETDKNKHGRNLNLYKPTTDRLLSGGTNLKGDITQTKG 610

Qy 99 ----NGRP 102
Db 611 KLFFSGRP 618

RESULT 12
D81019
adhesion and penetration protein NMB1985 [imported] - Neisseria meningitidis (group B strain MD58)
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 31-Dec-2000
C:Accession: D81019
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Qi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizzia, M. Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Venter, A.; Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Qi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizzia, M. Science 287, 1809-1815, 2000
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MD58.
A:Reference number: A81000; MUID:20175755
A:Accession: D81019
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1457 <TET>
A:Cross-references: GB:AE002547; GB:AE002098; NID:97227244; PIDN:AAF42312.1; PID:9722726
A:Experimental source: serogroup B, strain MMD58
C:Genetics:
A:Gene: NMB1985
C:Superfamily: Iga-specific metalloendopeptidase

Query Match 15.9%; Score 90.5; DB 2; Length 1457;
Best Local Similarity 25.6%; Pred. No. 0.37;
Matches 32; Conservative 20; Mismatches 38; Indels 35; Gaps 7;

Qy 8 YYALKSGRLNAPMPENGVA-----ENND--WVFMGYTOE-----ARKNAMNNKNN 52
Db 527 YFGFR--GGRLLD----NGHSLSFHRIQNTDEGAMIVNHNQDKESTVITGNKRIATTGNN 581

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 15:00:20 ; Search time 43.98 seconds
(without alignments)
162.510 Million cell updates/sec

Title: US-09-142-970-5
Perfect score: 565
Sequence: 1 LYKKNRYALKSGRLNAP.....NRFLTGKANLNGKISVTQ 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 198801 seqs, 68722935 residues

Total number of hits satisfying chosen parameters: 198801

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_67:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	559	98.9	1532	2 A26039	IgA-specific metal
2	494	87.4	1561	2 S61314	IgA-specific metal
3	494	87.4	1773	2 A81937	IgA-specific metal
4	490	86.7	1815	2 C81169	IgA-specific metal
5	298.5	52.8	1541	2 A37023	IgA-specific metal
6	297.5	52.7	1694	2 H64106	IgA-specific metal
7	297.5	52.7	1702	2 A41859	IgA-specific metal
8	291.5	51.6	1545	2 B41859	IgA-specific metal
9	287.5	50.9	1849	2 C41859	IgA-specific metal
10	132	23.4	1431	2 A81018	serine-type peptid
11	101	17.9	1394	2 S60762	IgA-specific serin
12	99.5	17.6	1457	2 D81019	adhesion and penet
13	98.5	17.4	709	2 C64057	IgA-specific metal
14	94.5	16.7	1449	2 B81963	IgA-specific serin
15	76	13.5	629	2 B75330	probable ribosomal
16	75	13.3	461	2 T10265	arabinogalactan-pr
17	74.5	13.2	451	2 A23535	clustered asparagi
18	74	13.1	163	2 T27368	hypothetical prote
19	74	13.1	210	2 F81386	probable acidic pe
20	73.5	13.0	1596	2 A33106	neurogenic locus m
21	72	12.7	419	2 T18420	hypothetical prote
22	71	12.6	232	2 T10646	hypothetical prote
23	70.5	12.5	1029	2 S64731	serine/threonine-s
24	70	12.4	1577	2 T30858	glucosyltransferas
25	69.5	12.3	599	2 A57701	sterol esterase (E
26	69.5	12.3	641	2 S57236	forked protein 2.5
27	69.5	12.3	958	2 A82583	conserved hypothet
28	69.5	12.3	1436	2 S57238	forked protein 5.4
29	69.5	12.3	1449	2 S57237	forked protein 5.6

30	68.5	12.1	402	2 B82418	hypothetical prote
31	68.5	12.1	406	2 D64934	hypothetical prote
32	68.5	12.1	604	2 S39885	forked protein - f
33	68.5	12.1	1247	2 E71616	hypothetical prote
34	68	12.0	549	2 T20720	hypothetical prote
35	68	12.0	631	2 S70908	transferrin-bindin
36	68	12.0	643	2 S28293	hypothetical prote
37	68	12.0	730	2 S28294	hypothetical prote
38	68	12.0	1336	2 T18288	hypothetical prote
39	68	12.0	1651	2 JCI340	ABC transport prot
40	68	12.0	2329	2 B64635	outer membrane pro
41	67.5	11.9	465	2 S41644	toxin-like outer m
42	67.5	11.9	537	2 A23770	polyadenylate-bind
43	67.5	11.9	612	2 A34967	asparagine-rich pr
44	67.5	11.9	642	2 D81401	sterol esterase (E
45	67.5	11.9	764	2 H71607	probable flagellar
					hypothetical prote

ALIGNMENTS

RESULT 1
A26039
IgA-specific metalloendopeptidase (EC 3.4.24.13) precursor - Neisseria gonorrhoeae (s
N:Alternate names: IgA protease; immunoglobulin A1 proteinase
C:Species: Neisseria gonorrhoeae
A:Variety: strain MS11
C:Date: 05-Oct-1988 #sequence_revision_05-Oct-1988 #text_change 08-Dec-2000
S:Accession: A26039; S09386
R:Pohlner, J.; Halter, R.; Beyreuther, K.; Meyer, T.F.
Nature 325, 458-462, 1987
A:Title: Gene structure and extracellular secretion of Neisseria gonorrhoeae IgA prot
A:Reference number: A26039; MUID:87115823
A:Accession: A26039
A:Molecule type: DNA
A:Residues: 1-1532 <POH>
A:Cross-references: GB:X04835; NID:944868; PIDN:CAA28538.1; PID:944869
A:Note: the authors translated the codon AAG for residue 668 as Asn
R:Halter, R.; Pohlner, J.; Meyer, T.F.
EMBO J. 8, 2737-2744, 1989
A:Title: Mosaic-like organization of IgA protease genes in Neisseria gonorrhoeae gene
A:Reference number: S09386; MUID:90060036
A:Accession: S09386
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 281-325, 'N', 327-337, 'N', 339-427, 'M', 429-531, 'N', 533-615, 'V', 617-631, 'N', 6
A:Experimental source: strain MS11
C:Genetics:
A:Gene: IgA
C:Superfamily: IgA-specific metalloendopeptidase
C:Keywords: hydrolase; metalloproteinase; transmembrane protein
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-1532/Product: immunoglobulin A1 proteinase #status predicted <MAT>
F:986-987/Cleavage site: Pro-Ser (autolytic) #status predicted
F:1018-1019/Cleavage site: Pro-Ser (autolytic) #status predicted
F:1121-1122/Cleavage site: Pro-Ala (autolytic) #status predicted

Query Match 98.9%; Score 559; DB 2; Length 1532;
Best Local Similarity 99.0%; Pred. No. 6.1e-50;
Matches 103; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	LYKKNRYALKSGRLNAPMPNGVAENNDWIFMGVTQEARKNMHNRRIGDFGG 60
DB	584	LYKKNRYALKSGRLNAPMPNGVAENNDWIFMGVTQEARKNMHNRRIGDFGG 643
QY	61	FFDEENGKHGALNLFNFGSKSQNRFLLTGGANLNGKISVTQ 104
DB	644	FFDEENGKHGALNLFNFGSKSQNRFLLTGGANLNGKISVTQ 687

RESULT 2
S61314

	Query Match	52.7%	Score 297.5;	DB 2;	Length 1694;
	Best Local Similarity	51.9%;	Pred. No. 1e-22;		
	Matches	54;	Conservative 17;	Mismatches 32;	Indels 1; Gaps 1;
Qy	1	LYYKNYYAYALKSGGRUNAPMPENGV	VAENNDWIPMGYTQEEARKNMMHNRRIGDFGG	60	
Db	601	LNLENYYYALRKGASTRSELPKNSGS	ENWLYMGKTSDEAKRNVNMHNINRMNGFNG	660	
Qy	61	PFDEENGKHNGALNLFNGKSAQNRELLT	GGANLNGKISVTOG	104	
Db	661	YFGEEEGK-NNGNLNVYFKGSEONRELLT	GGTNLNGDLKVEKG	703	

Query Match	51.6%	Score	291.5	DB	2	Length	1545
Best Local Similarity	52.8%	Pred. No.	4e-22				
Matches	56	Conservative	16	Mismatches	31	Indels	3
Gaps	2						
Qy	1	LY--KNRYALKSGRLNAPMPENGVAEDNWIPIFGYTQEEARKNAMNHKNRRIGDF	58				
		: : : : : : :					
Db	597	LYFNEENRYTALKDASIRSEFPQNRGESNNWLYMGTEKADAKQNMHNINRMNGF	656				
Qy	59	GGFFDEENGKHGALNMLNFGKSAQNRFLLTGCGANLNGKISVTQG	104				
		: : :					
Db	657	NGYFGEPEGR--NNGNLNVTEKFGKSEONRFLLTGTNTLNGDLNVOOG	701				

RESULT 9

[illegible]

B75330

Search completed: May 1, 2001, 15:00:21
Job time: 588 sec

THIS PAGE BLANK (USPTO)

Result No.	Query			ID	Description
	Score	Match	Length		
1	518	91.7	996	2	O30574 neisseria m
2	514	91.0	993	2	O9S6X3 neisseria m
3	514	91.0	997	2	O9S6X2 neisseria m
4	501	88.7	996	2	O57309 neisseria m
5	494	87.4	992	2	O30573 neisseria m
6	494	87.4	992	2	O57035 neisseria m
7	494	87.4	992	2	O9S6X5 neisseria m
8	494	87.4	992	2	O9S6X4 neisseria m
9	494	87.4	997	2	O30575 neisseria m
10	494	87.4	1561	2	O31169 neisseria m
11	490	86.7	1773	2	O9SVB9 neisseria m
12	490	86.7	1815	2	O9K0B4 neisseria m
13	132	23.4	1431	2	O9JXX3 neisseria m
14	99.5	17.6	1457	2	O9X7H1 neisseria m
15	99.5	17.6	1457	2	O9JXL6 neisseria m
16	94.5	16.7	1449	2	O9TWE4 neisseria m
17	78	13.8	277	13	O9PUY9 leplisosteus
18	76	13.5	629	2	O9SY6 deinococcus
19	75.5	13.4	943	10	O9S1T1 arabidopsis

KNOWN 1.
not over

KW	FT	FT	FT	SO
PLOCASE.				
	NON_TER	NON_TER	SEQUENCE	
	1	997	997 AA:	109811 MW:
	1	997		06F2E361E7E202E0 CRC64:

```

AC O30573;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98010345; PubMed=9350862;
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J.-F.,
RA del Valle J., Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012205; AAC45786.1; -
DR INTERPRO; IPR000710; -
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT SEQUENCE 992 AA; 109255 MW; EF7CA563BC7D7ECB CRC64;
SQ

Query Match 87.4%; Score 494; DB 2; Length 992;
Best Local Similarity 87.5%; Pred. No. le-43;
Matches 91; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGGRNAPMPENGVAENNDWIFMGVTOEARKNANMHNKNNRRIGDFGG 60
Db 557 LYKKNRYRYALKSGGSVNAPMPENGQTNNDWILMGSTQEEAKKNANMHNKNNRRISGFSG 616

QY 61 FFDEENGKGHGALNLFNGKSAQNRFLTLTGANLNGKISVTQG 104
Db 617 FFGEENGKGHGALNLFNGKSAQNRFLTLTGATNLNGKISVTQG 660

RESULT 6
Q57035 ID Q57035 PRELIMINARY; PRT; 992 AA.
AC Q57035;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98010345; PubMed=9350862;
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J.-F.,
RA del Valle J., Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012211; AAC45794.1; -
DR EMBL; X82480; CAA57863.1; -
DR EMBL; X82475; CAA57858.1; -

```

```

DR INTERPRO; IPR000710; -
DR INTERPRO; IPR002195; -
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT SEQUENCE 992 AA; 109228 MW; 3677DDE4CE6D9F69 CRC64;
SQ

Query Match 87.4%; Score 494; DB 2; Length 992;
Best Local Similarity 87.5%; Pred. No. le-43;
Matches 91; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGGRNAPMPENGVAENNDWIFMGVTOEARKNANMHNKNNRRIGDFGG 60
Db 557 LYKKNRYRYALKSGGSVNAPMPENGQTNNDWILMGSTQEEAKKNANMHNKNNRRISGFSG 616

QY 61 FFDEENGKGHGALNLFNGKSAQNRFLTLTGANLNGKISVTQG 104
Db 617 FFGEENGKGHGALNLFNGKSAQNRFLTLTGATNLNGKISVTQG 660

RESULT 7
Q9S6X5 ID Q9S6X5 PRELIMINARY; PRT; 992 AA.
AC Q9S6X5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J., del Valle J.,
RA Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012203; AAC45786.1; -
DR INTERPRO; IPR000710; -
DR INTERPRO; IPR002195; -
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT SEQUENCE 992 AA; 109239 MW; 306FF451B7213D49 CRC64;
SQ

Query Match 87.4%; Score 494; DB 2; Length 992;
Best Local Similarity 87.5%; Pred. No. le-43;
Matches 91; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGGRNAPMPENGVAENNDWIFMGVTOEARKNANMHNKNNRRIGDFGG 60
Db 557 LYKKNRYRYALKSGGSVNAPMPENGQTNNDWILMGSTQEEAKKNANMHNKNNRRISGFSG 616

QY 61 FFDEENGKGHGALNLFNGKSAQNRFLTLTGANLNGKISVTQG 104
Db 617 FFGEENGKGHGALNLFNGKSAQNRFLTLTGATNLNGKISVTQG 660

RESULT 8
Q9S6X4 ID Q9S6X4 PRELIMINARY; PRT; 992 AA.
AC Q9S6X4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

```

DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN (1)
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J., del Valle J.,
RA Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012204; AAC45787.1; -
DR INTERPRO; IPR000710; -
DR INTERPRO; IPR002195; -
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1 992
FT SEQUENCE 992 AA; 109239 MW; 9ED0A41019917CE3 CRC64;
SQ

Query Match 87.4%; Score 494; DB 2; Length 992;
Best Local Similarity 87.5%; Pred. No. 1e-43;
Matches 91; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 LYKKNRYALKSGGSLNPNFNGKSAQNRFLLTGGANLNGKISVTQ 104
|||||
DB 557 LYKKNRYALKSGGSLNPNFNGKSAQNRFLLTGGANLNGKISVTQ 616

QY 61 PFDEENGKNGALNPNFNGKSAQNRFLLTGGANLNGKISVTQ 104
|||||
DB 617 PFGEENGKNGALNPNFNGKSAQNRFLLTGGANLNGKISVTQ 660

RESULT 9
ID Q30575 PRELIMINARY; PRT; 997 AA.
AC Q30575;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN (1)
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J.,
RA del Valle J., Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012210; AAC45793.1; -
DR INTERPRO; IPR000710; -
DR INTERPRO; IPR002195; -
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1 997
FT SEQUENCE 997 AA; 110152 MW; 4DE40DE594E5E28E CRC64;
SQ

Query Match 87.4%; Score 494; DB 2; Length 997;
Best Local Similarity 87.5%; Pred. No. 1e-43;
Matches 91; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 LYKKNRYALKSGGSLNPNFNGKSAQNRFLLTGGANLNGKISVTQ 60
|||||

DB 558 LYKKNRYALKSGGSLNPNFNGKSAQNRFLLTGGANLNGKISVTQ 617

QY 61 PFDEENGKNGALNPNFNGKSAQNRFLLTGGANLNGKISVTQ 104
|||||

DB 618 PFGEENGKNGALNPNFNGKSAQNRFLLTGGANLNGKISVTQ 661

RESULT 10
ID Q51169 PRELIMINARY; PRT; 1561 AA.
AC Q51169;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE IGAL PROTEASE.
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=HF13;
RX MEDLINE=953402961; PubMed=7783620;
RA Lomholt B., Poulsen K., Mogens K.;
RT "Comparative characterization of the iga gene encoding IgA1 protease
RT in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
RT influenzae";
RL Mol. Microbiol. 15:495-506(1995).
DR EMBL; X82474; CAA57857.1; -
DR MEROPS; S06.001; -
DR INTERPRO; IPR000710; -
DR INTERPRO; IPR002195; -
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
SQ SEQUENCE 1561 AA; 171848 MW; 1C96E291A00017D5 CRC64;

Query Match 87.4%; Score 494; DB 2; Length 1561;
Best Local Similarity 87.5%; Pred. No. 1.7e-43;
Matches 91; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 LYKKNRYALKSGGSLNPNFNGKSAQNRFLLTGGANLNGKISVTQ 60
|||||

DB 584 LYKKNRYALKSGGSLNPNFNGKSAQNRFLLTGGANLNGKISVTQ 643

QY 61 PFDEENGKNGALNPNFNGKSAQNRFLLTGGANLNGKISVTQ 104
|||||

DB 644 PFGEENGKNGALNPNFNGKSAQNRFLLTGGANLNGKISVTQ 687

RESULT 11
ID Q9JVB9 PRELIMINARY; PRT; 1773 AA.
AC Q9JVB9;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE IGAL PROTEASE (EC 3.4.21.7).
GN IGA OR NMA0905.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis N.P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;

RT "Complete DNA sequence of a serogroup A strain of Neisseria
 RL meningitidis 22491."
 DR Nature 404:502-506(2000).
 KW EMBL; AL162754; CAB84182.1; -.
 SQ PROtease; Hydrolase.
 SEQUENCE 1773 AA; 196350 MW; CAC19E7313D76CE1 CRC64;

Query Match 87.4%; Score 494; DB 2; Length 1773;
 Best Local Similarity 87.5%; Pred. No. 2e-43;
 Matches 91; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
 QY 1 LYKKNRYALKSGRLNAPMPENGVAENNDWIFMGYTQEEARKNAMHNKRRIGDFGG 60
 Db 573 LYKKNRYALKSGGVSNAPEPENGQENNDWILMGSTQEEARKNAMHNKRRIGDFSG 632
 QY 61 FFDENGKGHGNALNPNFGSKSAQNRELLTGGANLNGKISVTQ 104
 Db 633 FFGENGKGHGNALNPNFGSKSAQNRELLTGGTNLNGKISVTQ 676

RESULT 12
 Q9K0B4 ID Q9K0B4 PRELIMINARY; PRT; 1815 AA.
 AC Q9K0B4;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DE 01-OCT-2000 (Tremblrel. 15, Last annotation update)
 DE IGA-SPECIFIC SERINE ENDOPEPTIDASE.
 GN NMB0700.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / SEROGROUP B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tattelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Clecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58."
 RL Science 287:1809-1815(2000).
 DR EMBL; AE002424; AAF41117.1;
 DR TIGR; NMB0700; -.
 SQ SEQUENCE 1815 AA; 201077 MW; 2259D4D71762C57F CRC64;

Query Match 86.7%; Score 490; DB 2; Length 1815;
 Best Local Similarity 86.5%; Pred. No. 5.4e-43;
 Matches 90; Conservative 4; Mismatches 10; Indels 0; Gaps 0;
 QY 1 LYKKNRYALKSGRLNAPMPENGVAENNDWIFMGYTQEEARKNAMHNKRRIGDFGG 60
 Db 573 LYKKNRYALKSGGVSNAPEPENGQENNDWILMGSTQEEARKNAMHNKRRIGDFSG 632
 QY 61 FFDENGKGHGNALNPNFGSKSAQNRELLTGGANLNGKISVTQ 104
 Db 633 FFGENGKGHGNALNPNFGSKSAQNRELLTGGTNLNGKISVTQ 676

RESULT 13
 Q9JXK3 ID Q9JXK3 PRELIMINARY; PRT; 1431 AA.
 AC Q9JXK3;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DE 01-OCT-2000 (Tremblrel. 15, Last annotation update)

DE SERINE-TYPE PEPTIDASE.
 GN NMB1998.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MC58 / SEROGROUP B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tattelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Clecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58."
 RL Science 287:1809-1815(2000).
 DR EMBL; AE002549; AAF42325.1; -.
 DR TIGR; NMB1998; -.
 SQ SEQUENCE 1431 AA; 157645 MW; 6801F0E209C24EFC CRC64;

Query Match 23.4%; Score 132; DB 2; Length 1431;
 Best Local Similarity 27.0%; Pred. No. 2e-05;
 Matches 37; Conservative 21; Mismatches 43; Indels 36; Gaps 5;
 QY 3 YKNYR--YYALKSGRLNAPMPENGVAENNDWIFMGYTQEEARKNAMHNKRRIGDFG 59
 Db 555 HRNRRTDYFLKPGGNPREFFPLN-MKNSTSWQFIGNRQQAABQVAENRDLITFG 613
 QY 60 GFPDE--ENCKG-----HNGALNPNFGSKSAQNRF 87
 Db 614 GYLGNAQTGKAAPSYKTNDAIEKTRHIANAAYGRPEYRNGALNLHYRPRKTDSTL 673.
 QY 88 LLTGGANLNGKISVTQ 104
 Db 674 LLNGMNLNGEVLEGG 690

RESULT 14
 Q9X7H1 ID Q9X7H1 PRELIMINARY; PRT; 1457 AA.
 AC Q9X7H1;
 DT 01-NOV-1999 (Tremblrel. 12, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DE 01-JUN-2000 (Tremblrel. 14, Last annotation update)
 DE APP PROTEIN.
 GN APP.
 OS Neisseria meningitidis serogroup B.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B-15;P1.16;
 RA Abdel-Hadi H., Wooldridge K.G., Ala Aldeen D.A.;
 RT "Identification and characterisation of neisserial App."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ242535; CAB43832.2; -.
 DR INTERPRO; IPR000710; -.
 DR INTERPRO; IPR001254; -.
 DR PRINTS; PR00921; IGASERPTASE.
 DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
 SQ SEQUENCE 1457 AA; 159956 MW; CE62390B6C04B781 CRC64;

Query Match 17.6%; Score 99.5; DB 2; Length 1457;
 Best Local Similarity 27.4%; Pred. No. 0.054;
 Matches 32; Conservative 18; Mismatches 42; Indels 25; Gaps 4;
 QY 8 YYALKSGRLNAPMPENGVA-----ENNDWIFMGYTQEEARKNAMHNKRRIGDFG-- 59

```
Db 527 YFGFR-GGRDL-NGHSLSFHRIQNTDEGAMIVNHNDQKESVTITGNKDIATTGNN 581
Qy 60 -----GPFDEENGKGHNGALNLFNGKSAQNRFLLTGGANLNGKISVTQG 104
Db 582 NSLDSKKEIAYNGWFGKEDTTKTNGRLNLVYQPAADRTLLLSGGTNLNGNITQTNG 638

RESULT 15
QJXL6
ID QJXL6 PRELIMINARY; PRT: 1457 AA.
AC QJXL6;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE ADHESION AND PENETRATION PROTEIN.
CN NMB1985.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,
RA Cotton M.D., Otterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58."
RL Science 287:1809-1815(2000).
DR EMBL; AE002547; AAF42312.1; -.
DR TIGR; NMB1985; -.
SQ SEQUENCE 1457 AA; 159965 MW; B759529CFD4BD0AF CRC64;
```

```
Query Match 17.6%; Score 99.5; DB 2; Length 1457;
Best local Similarity 27.4%; Pred. No. 0.054;
Matches 32; Conservative 18; Mismatches 42; Indels 25; Gaps 4;

Qy 8 YYALKSGGRNLNAMPENGVA-----ENNDFMGYTOEARKNAMNNRRRIGDFG-- 59
Db 527 YFGFR-GGRDL-NGHSLSFHRIQNTDEGAMIVNHNDQKESVTITGNKDIATTGNN 581
Qy 60 -----GPFDEENGKGHNGALNLFNGKSAQNRFLLTGGANLNGKISVTQG 104
Db 582 NSLDSKKEIAYNGWFGKEDTTKTNGRLNLVYQPAADRTLLLSGGTNLNGNITQTNG 638
```

Search completed: May 1, 2001, 15:04:56
Job time: 322 sec


```

FT PROPEP      987 1532      HELPER PEPTIDE.
FT ACT_SITE    278 278      POTENTIAL.
FT SITE        986 987      CLEAVAGE (AUTO-).
FT SITE        1018 1019     CLEAVAGE (AUTO-).
FT SITE        1121 1122     CLEAVAGE (AUTO-).
SQ SEQUENCE    1532 AA; 168976 MW; 68FF4112BD22F40D CRC64;

Query Match      98.9%; Score 559; DB 1; Length 1532;
Best Local Similarity 99.0%; Pred. No. 2.6e-50;
Matches 103; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LYKKNRYALKSGGRLNAPMPENGVAENNDWIFMGYTQEEARKNMHNRRIGDFFG 60
   |||||
Db 584 LYKKNRYALKSGGRLNAPMPENGVAENNDWIFMGYTQEEARKNMHNRRIGDFFG 643
   |||||

QY 61 FDEENGKGHGKALNLFNGKSAQNRFLLTGANLNGKISVTQ 104
   |||||
Db 644 FDEENGKGHGKALNLFNGKSAQNRFLLTGANLNGKISVTQ 687
   |||||

RESULT 2
ID IGA0_HAEIN STANDARD; PRT; 1541 AA.
AC P42782;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 01-NOV-1995 (Rel. 32, Last annotation update)
DE IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (EC 3.4.21.72) (IGAI PROTEASE).
GN IGA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HK368 / SEROTYPE B;
RX MEDLINE=89379374; PubMed=2506130;
RA Poulsen K., Reinholdt J., Hjørth J.P., Thøgersen H.C., Killian M.;
RT "Cloning and sequencing of the immunoglobulin A1 protease gene (iga)
RL of Haemophilus influenzae serotype b.";
RN Infect. Immun. 57:3097-3105(1989).
RN [2]
RP MUTAGENESIS OF SER-288.
RC STRAIN-HK368 / SEROTYPE B;
RX MEDLINE=92234949; PubMed=1373717;
RA Poulsen K., Reinholdt J., Killian M.;
RT "A comparative genetic study of serologically distinct Haemophilus
RL influenzae type 1 immunoglobulin A1 proteases.";
CC -!- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT
CC CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE
CC SUBSTRATES ARE KNOWN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X64357; CAA45708.1;
DR EMBL; M87492; AAA24969.1;

```

```

DR MEROPS; S06.001; -.
DR INTERPRO; IPR000710; -.
DR PRINTS; PR00921; IGASERPTASE.
KW Hydrolase; Serine protease; Transmembrane; Zymogen; Signal.
FT SIGNAL      1 25      POTENTIAL.
FT CHAIN       26 1008     IMMUNOGLOBULIN A1 PROTEASE.
FT PROPEP     1009 1541     HELPER PEPTIDE (POTENTIAL).
FT ACT_SITE   288 288     PROBABLE.
FT MUTAGEN    288 288     S->T: LOSS OF ACTIVITY.
SQ SEQUENCE   1541 AA; 169370 MW; CE7257CB3196C600 CRC64;

Query Match      52.8%; Score 298.5; DB 1; Length 1541;
Best Local Similarity 51.9%; Pred. No. 3.3e-23;
Matches 54; Conservative 18; Mismatches 31; Indels 1; Gaps 1;

QY 1 LYKKNRYALKSGGRLNAPMPENGVAENNDWIFMGYTQEEARKNMHNRRIGDFFG 60
   |||||
Db 595 LNLENTYALRKGASTRSELPKSGESNENWLYMGKTSDEAKRVNMHNRRINRMNGFNG 654
   |||||

QY 61 FDEENGKGHGKALNLFNGKSAQNRFLLTGANLNGKISVTQ 104
   |||||
Db 655 YFGEEGK-NNGNLNVTFGKSEQNRFLLTGTLNGDLTVK 697
   |||||

RESULT 3
ID IGA0_HAEIN STANDARD; PRT; 1694 AA.
AC P44969;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (EC 3.4.21.72) (IGAI PROTEASE).
GN IGA OR IGA1 OR HI0990.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SEROTYPE D;
RA Wright A., Fishman Y., Tai F., Plaut A.G.;
RL Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McInerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
CC PRODUCING INTACT FC AND FAB FRAGMENTS.
CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT
CC CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE
CC SUBSTRATES ARE KNOWN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
CC SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```


CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X59800; ; NOT_ANNOTATED_CDS.

DR EMBL; U32779; AAC22651.1; ;

DR MEROPS; S06.001; ;

DR TIGR; H10990; ;

DR InterPro; IPR000710; ;

DR PRINTS; PR00921; IGASERPTASE.

KW Hydrolase; Serine protease; Transmembrane; Zymogen; Signal.

FT SIGNAL 1 25 POTENTIAL.

FT CHAIN 26 1014 IMMUNOGLOBULIN A1 PROTEASE.

FT PROPEP 1015 1694 HELPER PEPTIDE (POTENTIAL).

FT ACT_SITE 288 288 PROBABLE.

FT CONFLICT 253 254 EN -> GV (IN REF. 1).

FT CONFLICT 272 272 G -> A (IN REF. 1).

FT CONFLICT 464 464 G -> E (IN REF. 1).

FT CONFLICT 866 866 S -> T (IN REF. 1).

FT CONFLICT 1036 1036 A -> D (IN REF. 1).

FT CONFLICT 1074 1074 A -> G (IN REF. 1).

FT CONFLICT 1421 1421 A -> G (IN REF. 1).

FT CONFLICT 1545 1545 H -> T (IN REF. 1).

SQ SEQUENCE 1694 AA; 185539 MW; C52427013F93178C CRC64;

Query Match 52.78; Score 297.5; DB 1; Length 1694;

Best Local Similarity 51.98; Pred. No. 4.6e-23;

Matches 54; Conservative 17; Mismatches 32; Indels 1; Gaps 1;

OY 1 LYKNRYVALKSGRLNAPNGVAENNDWIFMGVTOEARKNANMKNRNRIGDFGG 60

Db 601 LNLNRYVALKSGRLNAPNGVAENNDWIFMGVTOEARKNANMKNRNRIGDFGG 60

OY 61 FFDENGKHGALNLFNGKSAQNRELLTGGANLNGKISVTQ 104

Db 661 YFGEEGK-NNGNLNVTFKKGSEQNRELLTGGTNGDLKVEKG 703

RESULT 4

IGA2_HAEIN

ID IGA2_HAEIN STANDARD; PRT; 1702 AA.

AC P45384;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (EC 3.4.21.72) (IGAL PROTEASE).

GN IGA.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Haemophilus.

OX NCBI_TaxID=727;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HK15 / SEROTYPE B;

RX MEDLINE=92234949; PubMed=1373717;

RT Poulsen K., Reinholdt J., Kilian M.;

RT "A comparative genetic study of serologically distinct Haemophilus

influenzae type 1 immunoglobulin A1 proteases.";

RL J. Bacteriol. 174:2913-2921(1992).

CC -1- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A

CC PRODUCING INTACT FC AND FAB FRAGMENTS.

CC CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT

CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE

CC SUBSTRATES ARE KNOWN.

CC -1- SUBCELLULAR LOCATION: SECRETED.

CC -1- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC

SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE

OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE

CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY

CC SIMILARITY).

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M87489; AAA24966.1; ;

DR MEROPS; S06.001; ;

DR InterPro; IPR000710; ;

DR PRINTS; PR00921; IGASERPTASE.

KW Hydrolase; Serine protease; Transmembrane; Zymogen; Signal.

FT SIGNAL 1 25 POTENTIAL.

FT CHAIN 26 1014 IMMUNOGLOBULIN A1 PROTEASE.

FT PROPEP 1015 1702 HELPER PEPTIDE (POTENTIAL).

FT ACT_SITE 288 288 PROBABLE.

FT DOMAIN 1109 1124 2 X 8 AA TANDEM REPEATS OF A-K-V-E-K-E-E-

FT REPEAT 1109 1116 1.

FT REPEAT 1117 1124 2.

SQ SEQUENCE 1702 AA; 186539 MW; 860F70D2667807A6 CRC64;

Query Match 52.78; Score 297.5; DB 1; Length 1702;

Best Local Similarity 51.98; Pred. No. 4.6e-23;

Matches 54; Conservative 17; Mismatches 32; Indels 1; Gaps 1;

OY 1 LYKNRYVALKSGRLNAPNGVAENNDWIFMGVTOEARKNANMKNRNRIGDFGG 60

Db 601 LNLNRYVALKSGRLNAPNGVAENNDWIFMGVTOEARKNANMKNRNRIGDFGG 60

OY 61 FFDENGKHGALNLFNGKSAQNRELLTGGANLNGKISVTQ 104

Db 661 YFGEEGK-NNGNLNVTFKKGSEQNRELLTGGTNGDLKVEKG 703

RESULT 5

IGA3_HAEIN

ID IGA3_HAEIN STANDARD; PRT; 1545 AA.

AC P45385;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (EC 3.4.21.72) (IGAL PROTEASE).

GN IGA.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Haemophilus.

OX NCBI_TaxID=727;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HK393 / NCTC 8467 / SEROTYPE B;

RX MEDLINE=92234949; PubMed=1373717;

RT Poulsen K., Reinholdt J., Kilian M.;

RT "A comparative genetic study of serologically distinct Haemophilus

influenzae type 1 immunoglobulin A1 proteases.";

RL J. Bacteriol. 174:2913-2921(1992).

CC -1- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A

CC PRODUCING INTACT FC AND FAB FRAGMENTS.

CC CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT

CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE

CC SUBSTRATES ARE KNOWN.

CC -1- SUBCELLULAR LOCATION: SECRETED.

CC -1- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC

SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE

OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE

CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY

CC SIMILARITY).

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).

[illegible]

Db 925 FLFNNSSSTNSSECS 942

RESULT 11

ID	MAM_DROME	STANDARD;	PRT;	1596 AA.
AC	F21519;			
DT	01-MAY-1991 (Rel. 18, Created)			
DT	01-MAY-1991 (Rel. 18, Last sequence update)			
DT	01-MAR-1992 (Rel. 21, Last annotation update)			
DE	NEUROGENIC PROTEIN MASTERMIND.			
GN	MAM.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Eukaryota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
ON	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CANTON-S;			
RX	MEDLINE=91065516; PubMed=1701150;			
RA	Smoller D., Friedel C., Schmid A., Bettler D., Lam L., redvobnick B.;			
RT	"The Drosophila neurogenic locus mastermind encodes a nuclear protein unusually rich in amino acid homopolymers.";			
RL	Genes Dev. 4:1688-1700(1990).			
CC	-!- FUNCTION: MAY HAVE A REGULATORY FUNCTION POSSIBLY IN ASSOCIATION WITH THE N GENE PRODUCT.			
CC	-!- SUBCELLULAR LOCATION: NUCLEAR.			
CC	-!- DEVELOPMENTAL STAGE: DURING EARLY NEUROGENESIS MAM PRODUCTS ARE UBIQUITOUSLY LOCATED. DURING LATER STAGES THEY ACCUMULATE IN THE CENTRAL NERVOUS SYSTEM.			
CC	-!- MISCELLANEOUS: THE PROTEIN HAS MANY AA HOMOPOLYMERIC DOMAINS: 21 POLY-GLN RUNS (FROM 5 TO 16 AA IN LENGTH), 4 POLY-GLY (6 TO 10 AA), 3 POLY-ASN (3 X 5 AA), 1 POLY-ALA (10 AA) AND 1 POLY-THR (5 AA) RUNS.			
CC	-!- SIMILARITY: TO OTHER NUCLEAR PROTEINS OF DROSOPHILA, TO CERTAIN YEAST AND MAMMALIAN REGULATORY PROTEINS.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).			
CC	EMBL; X54251; CAA38152.1; -			
DR	PIR; A33106; A33106.			
DR	PIR; A36391; A36391.			
DR	FlyBase; Fgn0002643; mam.			
KW	Neurogenesis; Nuclear protein; Repeat.			
FT	DOMAIN 20 84 GLN-RICH.			
FT	DOMAIN 127 190 ARG/LYS-RICH (BASIC).			
FT	DOMAIN 196 219 GLN-RICH.			
FT	DOMAIN 259 304 ASN-RICH.			
FT	DOMAIN 355 388 GLY/ASN-RICH.			
FT	DOMAIN 392 406 GLN-RICH.			
FT	DOMAIN 407 440 GLY-RICH.			
FT	DOMAIN 651 671 GLN-RICH.			
FT	DOMAIN 700 714 GLN-RICH.			
FT	DOMAIN 759 816 GLN-RICH.			
FT	DOMAIN 987 996 5 X 2 AA TANDEM REPEATS OF G-V.			
FT	DOMAIN 1060 1079 ALA-RICH.			
FT	DOMAIN 1092 1107 8 X 2 AA TANDEM REPEATS OF V-G.			
FT	DOMAIN 1237 1252 7 X 2 AA TANDEM REPEATS OF G-V.			
FT	DOMAIN 1452 1496 POLY-THR.			
FT	DOMAIN 1599 1592 ASP/GLU-RICH (ACIDIC).			
FT	DOMAIN 1596 AA; 167717 MW; B944D86EF359D605 CRC64;			
SC	SEQUENCE			

Query Match 13.0%; Score 73.5; DB 1; Length 1596;
Best Local Similarity 35.1%; Pred. No. 8.6;

```

FT ZN_FING 496 520 C2H2-TYPE.
FT VARSPIC 55 148 MISSING (IN ISOFORM IK-2, ISOFORM IK-4
AND ISOFORM IK-8).
FT VARSPIC 55 291 MISSING (IN ISOFORM IK-6).
FT VARSPIC 149 291 MISSING (IN ISOFORM IK-5).
FT VARSPIC 206 246 MISSING (IN ISOFORM IK-4).
FT VARSPIC 206 291 MISSING (IN ISOFORM IK-3).
FT VARSPIC 247 291 MISSING (IN ISOFORM IK-6 AND ISOFORM IK-
8).
FT SQ SEQUENCE 522 AA; 57657 MW; EE9DA9B2A0ECF88A CRC64;

Query Match 12.6%; Score 71; DB 1; Length 522;
Best Local Similarity 31.1%; Pred. No. 4.5;
Matches 28; Conservative 12; Mismatches 28; Indels 22; Gaps 6;

QY 22 PENCVAENNDWIPMGVYQEEARKKAMNHKNNRRIGDGGF-----FDEENGKHNG 72
| | | | | : : : : | | | | | | | | | | | | | | |
Db 19 PPNDVSENDGA-MPIPEDLSASNNLGH-NNR--GDKEGLACNIKVEARCDENGL----- 70
| | | | | : : : : | | | | | | | | | | | | | | |

QY 73 ALNLFNGKSAQN-----RFLITGGANLNG 97
| | | | | : : : : | | | | | | | | | | | | | | |
Db 71 AIDMMNGEENEECAEDLRVLDSAGKVG 100
| | | | | : : : : | | | | | | | | | | | | | | |

RESULT 13
KSPI_YEAST STANDARD; PRT; 1029 AA.
AC P38691;
CD 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SRINE/THREONINE-PROTEIN KINASE KSPI (EC 2.7.1.-).
GN KSPI OR YHR082C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M335 /2A.
RX MEDLINE=96194625; PubMed=8676864;
RT Fieischmann M., Stagljar I., Aebi M.;
RT "Allele-specific suppression of a Saccharomyces cerevisiae prp20
RT mutation by overexpression of a nuclear serine/threonine protein
RT kinase.";
RL Mol. Gen. Genet. 250:614-625(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Latrelle P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
RA Nham M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaubin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII.";
RL Science 265:2077-2082(1994).
CC -!- FUNCTION: MAY ACT ON PRP20.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-
CC PROTEIN KINASES. BELONGS TO THE CK-II SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

DR HIV; M31325; VPR$SMMPBJ.
DR InterPro: IPR000012; -.
DR Pfam: PF00522; VPR; 1.
DR PRINTS; PR00444; HIVVPRVPX.
KW AIDS.
SQ SEQUENCE 101 AA; 11429 MW; 0D2156A07042EE98 CRC64;

Query Match      12.3%; Score 69.5; DB 1; Length 101;
Best Local Similarity 23.4%; Pred. No. 1;
Matches 22; Conservative 18; Mismatches 43; Indels 11; Gaps 2;

QY 22 PNCVAENNDW-IFMGYTOEARKNAMNNRRRIGDFGFFDEENGKGHG----- 72
||| : : : || : : : | : : : | : : : | : : : |
Db 6 PEDEAPQREPDWDVVLEVEIEEELAHNHPDLLTALGNIIYDRHGDTLLEGAGLIRIL 65
|| : : : : : || : : : | : : : | : : : | : : : |

QY 73 --ALNLFNGKSAQNRLFLTGANLNKISVTQG 104
|| : : : : : || : : : | : : : | : : : | : : : |
Db 66 QRALFTIHRGGCRHSRIGQSQQGNPLSTIPPSRG 99
|| : : : : : || : : : | : : : | : : : | : : : |

RESULT 15
ID BAL_MOUSE STANDARD; PRT; 599 AA.
BAL_MOUSE AC O64285;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE BILE-SALT-ACTIVATED LIPASE PRECURSOR (EC 3.1.1.3) (EC 3.1.1.13) (BAL)
DE (BILE-SALT-STIMULATED LIPASE) (BSSL) (CARBOXYL ESTER LIPASE) (STEROL
DE ESTERASE) (CHOLESTEROL ESTERASE) (PANCREATIC LYSPHOSPHOLIPASE).
DE CEL OR LIPI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Mammary gland;
RX MEDLINE=96096531; PubMed=8522186;
RA Mackay K., Lawn R.M.;
RT "Molecular cloning and characterization of the mouse carboxyl ester
RT lipase gene and evidence for expression in the lactating mammary
RT gland."
RL Genomics 29:115-122(1995).
CC -!- FUNCTION: CATALYZES FAT AND VITAMIN ABSORPTION. ACTS IN CONCERT
CC WITH PANCREATIC LIPASE AND COLIPASE FOR THE COMPLETE DIGESTION
CC OF DIETARY TRIGLYCERIDES (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: TRIACYLGlycerol + H(2)O = DIACYLGLYCEROL + A
CC FATTY ACID ANION.
CC -!- CATALYTIC ACTIVITY: A STERYL ESTER + H(2)O = A STEROL + A FATTY
CC ACID.
CC -!- SIMILARITY: BELONGS TO THE TYPE-B CARBOXYLESTERASE/LIPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U33169; AAA92088.1; -.
CC EMBL; U37386; AAC52279.1; -.
CC HSSP; P30122; LAQL.

```

Search completed: May 1, 2001, 15:03:39
Job time: 286 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 15:04:53 ; Search time 73.06 Seconds
(without alignments)
166.844 Million cell updates/sec

Title: US-09-142-970-4
Perfect score: 569
Sequence: 1 LYKKNRYRYALKSGRLNAP.....NRFLTTGGANLNGNGRPVK 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_15:*
- 1: sp_archaea:*
 - 2: sp_bacteria:*
 - 3: sp_fungi:*
 - 4: sp_human:*
 - 5: sp_invertebrate:*
 - 6: sp_mammal:*
 - 7: sp_mhc:*
 - 8: sp_organelle:*
 - 9: sp_phase:*
 - 10: sp_plant:*
 - 11: sp_rodent:*
 - 12: sp_unclassified:*
 - 13: sp_vertebrate:*
 - 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	484	85.1	996	2 O30574	O30574 neisseria m
2	480	84.4	993	2 Q9S6X3	Q9S6X3 neisseria m
3	480	84.4	997	2 Q9S6X2	Q9S6X2 neisseria m
4	467	82.1	996	2 O57309	O57309 neisseria m
5	458	80.5	992	2 O30573	O30573 neisseria m
6	458	80.5	992	2 Q57035	Q57035 neisseria m
7	458	80.5	992	2 Q9S6X5	Q9S6X5 neisseria m
8	458	80.5	992	2 Q9S6X4	Q9S6X4 neisseria m
9	458	80.5	997	2 O30575	O30575 neisseria m
10	458	80.5	1561	2 O51169	O51169 neisseria m
11	458	80.5	1773	2 Q9JVB9	Q9JVB9 neisseria m
12	454	79.8	1615	2 Q9K0B4	Q9K0B4 neisseria m
13	134	23.6	1431	2 Q9JXK3	Q9JXK3 neisseria m
14	90.5	15.9	1457	2 Q9X7H1	Q9X7H1 neisseria m
15	90.5	15.9	1457	2 Q9JXL6	Q9JXL6 neisseria m
16	87.5	15.4	319	10 Q41725	Q41725 zinnia eleg
17	84.5	14.9	1449	2 Q9JWB4	Q9JWB4 neisseria m
18	83	14.6	549	5 Q19318	Q19318 caenorhabdi
19	81.5	14.3	809	5 P90534	P90534 dictyosteli

20	80.5	14.1	586	5 Q9VGH5	Q9VGH5 drosophila
21	80	14.1	456	5 Q20936	Q20936 caenorhabdi
22	80	14.1	461	10 Q41256	Q41256 nicotiana a
23	79	13.9	1336	5 Q94479	Q94479 dictyosteli
24	78.5	13.8	419	5 O77316	O77316 plasmodium
25	78.5	13.8	807	5 O96262	O96262 plasmodium
26	78	13.7	568	5 Q9NL38	Q9NL38 pinctada ma
27	78	13.7	1449	2 Q55264	Q55264 streptococ
28	76.5	13.4	560	5 O61085	O61085 dictyosteli
29	76	13.4	163	5 Q9XWCL	Q9XWCL caenorhabdi
30	76	13.4	1212	5 Q9UOL0	Q9UOL0 plasmodium
31	76	13.4	1577	2 Q55265	Q55265 streptococ
32	75.5	13.3	263	2 Q9L1S4	Q9L1S4 streptomyce
33	75.5	13.3	949	5 O97306	O97306 plasmodium
34	75.5	13.3	1538	2 O53395	O53395 mycobacteri
35	75.5	13.3	2150	5 Q23863	Q23863 dictyosteli
36	75	13.2	317	4 O95024	O95024 homo sapien
37	75	13.2	1408	5 Q24341	Q24341 drosophila
38	74.5	13.1	631	2 Q48043	Q48043 haemophilus
39	74.5	13.1	1189	5 Q9VVR4	Q9VVR4 drosophila
40	74	13.0	277	13 Q9PUY9	Q9PUY9 lepisosteus
41	74	13.0	290	10 Q9L249	Q9L249 arabidopsi
42	74	13.0	472	2 O07121	O07121 lactococcus
43	73.5	12.9	319	2 Q9RF69	Q9RF69 neisseria g
44	73.5	12.9	319	2 Q9RF57	Q9RF57 neisseria g
45	73.5	12.9	348	2 P95346	P95346 neisseria g

ALIGNMENTS

RESULT 1
O30574 PRELIMINARY; PRT; 996 AA.
AC O30574
DT 01-JAN-1998 (TREMREL. 05, Created)
DT 01-JAN-1998 (TREMREL. 05, Last sequence update)
DT 01-OCT-2000 (TREMREL. 15, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98010345; PubMed=9350862;
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J.-F.,
RA del Valle J., Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread."
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL: AF012206; RAC45789;
DR INTERPRO: IPR000710; -;
DR INTERPRO: IPR002195; -;
DR PRINTS: PR00921; IGASERPTASE.
DR PROSITE: PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER
FT NON_TER
SQ SEQUENCE 996 AA; 109688 MW; B3ABDEF54C337D9 CRC64;

Query Match 85.1%; Score 484; DB 2; Length 996;
Best Local Similarity 82.1%; Pred. No. 1.4e-41;
Matches 92; Conservative 4; Mismatches 6; Indels 10; Gaps 2;

QY 1 LYKKNRYRYALKSGRLNAPMPNGVAENNDWYFMGYTQEEARKNNNNRRIGDFG 60
|||||
Db 557 LYKKNRYRYALKSGSVNAPMPNGVTENNNDWYFMGYTQEEAKNNNNRRIGSFG 616
|||||

QY 61 FDEENGKNGHNGALNLFNNGKSAQNRFLLTGGANLNG-----GN----GRP 102
|||
Db 617 FFEENGKNGHNGALNLFNNGKSAQNRFLLTGGANLNGKISVTQGNVLLSGRP 668
|||


```

AC O30573;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE IGAI PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98010345; PubMed=9350862;
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J.-F.,
RA del Valle J., Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012205; AAC45788.1; -
DR INTERPRO; IPR000710; -.
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 992
SQ SEQUENCE 992 AA; 109255 MW; EF7CA563BC7D7ECB CRC64;

Query Match 80.5%; Score 458; DB 2; Length 992;
Best Local Similarity 77.7%; Pred. No. 6.4e-39;
Matches 87; Conservative 5; Mismatches 10; Indels 10; Gaps 2;

Qy 1'LYKNNRYALKSGRLNAPNPENGVAENNDWFMGYTQEEARKNAMNKNRRIGDFGG 60
Db 557 LYKNNRYALKSGSYNAPNPENGQTNNDWILMGSTQEEAKNAMNKNQRISGFSG 616

Qy 61 FFEENGKGHGALNLFNGKSAQNRFLLTGGANLNG-----GN----GRP 102
Db 617 FFEENGKGHGALNLFNGKSAQNRFLLTGGTNLNGKISVVTQGNVLLSGRP 668

RESULT 6
ID Q57035 PRELIMINARY; PRT; 992 AA.
AC Q57035;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE IGAI PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98010345; PubMed=9350862;
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J.-F.,
RA del Valle J., Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012211; AAC45794.1; -
DR INTERPRO; IPR000710; -.
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 992
SQ SEQUENCE 992 AA; 109255 MW; EF7CA563BC7D7ECB CRC64;

Query Match 80.5%; Score 458; DB 2; Length 992;
Best Local Similarity 77.7%; Pred. No. 6.4e-39;
Matches 87; Conservative 5; Mismatches 10; Indels 10; Gaps 2;

Qy 1'LYKNNRYALKSGRLNAPNPENGVAENNDWFMGYTQEEARKNAMNKNRRIGDFGG 60
Db 557 LYKNNRYALKSGSYNAPNPENGQTNNDWILMGSTQEEAKNAMNKNQRISGFSG 616

Qy 61 FFEENGKGHGALNLFNGKSAQNRFLLTGGANLNG-----GN----GRP 102
Db 617 FFEENGKGHGALNLFNGKSAQNRFLLTGGTNLNGKISVVTQGNVLLSGRP 668

RESULT 7
ID Q9S6X5 PRELIMINARY; PRT; 992 AA.
AC Q9S6X5;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE IGAI PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J., del Valle J.,
RA Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012203; AAC45786.1; -
DR INTERPRO; IPR000710; -.
DR PRINTS; IPR002195; -
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 992
SQ SEQUENCE 992 AA; 109239 MW; 306FF451B7213D49 CRC64;

Query Match 80.5%; Score 458; DB 2; Length 992;
Best Local Similarity 77.7%; Pred. No. 6.4e-39;
Matches 87; Conservative 5; Mismatches 10; Indels 10; Gaps 2;

Qy 1 LYKNNRYALKSGRLNAPNPENGVAENNDWFMGYTQEEARKNAMNKNRRIGDFGG 60
Db 557 LYKNNRYALKSGSYNAPNPENGQTNNDWILMGSTQEEAKNAMNKNQRISGFSG 616

Qy 61 FFEENGKGHGALNLFNGKSAQNRFLLTGGANLNG-----GN----GRP 102
Db 617 FFEENGKGHGALNLFNGKSAQNRFLLTGGTNLNGKISVVTQGNVLLSGRP 668

RESULT 8
ID Q9S6X4 PRELIMINARY; PRT; 992 AA.
AC Q9S6X4;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TReMBLrel. 14, Last annotation update)

```

```

DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J., del Valle J.,
RA Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064 (1997).
DR EMBL; AF012204; AAC45787.1; -.
DR INTERPRO; IPR000710; -.
DR INTERPRO; IPR002195; -.
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 992
SQ SEQUENCE 992 AA; 109239 MW; 9ED0A41019917CE3 CRC64;

Query Match 80.5%; Score 458; DB 2; Length 992;
Best Local Similarity 77.7%; Pred. No. 6.4e-39;
Matches 87; Conservative 5; Mismatches 10; Indels 10; Gaps 2;

Qy 1 LYKKNRYALKSGGSLNPNPENGQTNNDWILMGSTQEEAKKNMKNRRIGDFGG 60
Db 557 LYKKNRYALKSGGSLNPNPENGQTNNDWILMGSTQEEAKKNMKNRRIGDFGG 616

Qy 61 PFDEENGKNGALNPNFNGKSAQNRFLLTGGANLNG-----GN-----GRP 102
Db 617 PFGEENGKNGALNPNFNGKSAQNRFLLTGGTNLNGKISVTQGNVLLSGRP 668

RESULT 9.
O30575 PRELIMINARY; PRT; 997 AA.
AC O30575;
DT 01-JAN-1998 (TRENBLrel. 05, Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J.-F.,
RA del Valle J., Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064 (1997).
DR EMBL; AF012210; AAC45793.1; -.
DR INTERPRO; IPR000710; -.
DR INTERPRO; IPR002195; -.
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 997
SQ SEQUENCE 997 AA; 110152 MW; 4DE40DE594E5E28E CRC64;

Query Match 80.5%; Score 458; DB 2; Length 997;
Best Local Similarity 77.7%; Pred. No. 6.4e-39;
Matches 87; Conservative 5; Mismatches 10; Indels 10; Gaps 2;

Qy 1 LYKKNRYALKSGGSLNPNPENGQTNNDWILMGSTQEEAKKNMKNRRIGDFGG 60

```

```

Db 558 LYKKNRYALKSGGSLNPNPENGQTNNDWILMGSTQEEAKKNMKNRRIGDFGG 617
Qy 61 PFDEENGKNGALNPNFNGKSAQNRFLLTGGANLNG-----GN-----GRP 102
Db 618 PFGEENGKNGALNPNFNGKSAQNRFLLTGGTNLNGKISVTQGNVLLSGRP 669

RESULT 10
O51169 PRELIMINARY; PRT; 1561 AA.
AC O51169;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DE 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE IGAL PROTEASE.
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=HF13;
RX MEDLINE=95302961; PubMed=7783620;
RA Lomholt H., Poulsen K., Mogens K.;
RT "Comparative characterization of the iga gene encoding IgA1 protease
RT in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
RT influenzae.";
RL Mol. Microbiol. 15:495-506 (1995).
DR EMBL; X82474; CAA57857.1; -.
DR MEROPS; S06.001; -.
DR INTERPRO; IPR000710; -.
DR INTERPRO; IPR002195; -.
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
SQ SEQUENCE 1561 AA; 171848 MW; 1C96E291A00017D5 CRC64;

Query Match 80.5%; Score 458; DB 2; Length 1561;
Best Local Similarity 77.7%; Pred. No. 1.1e-38;
Matches 87; Conservative 5; Mismatches 10; Indels 10; Gaps 2;

Qy 1 LYKKNRYALKSGGSLNPNPENGQTNNDWILMGSTQEEAKKNMKNRRIGDFGG 60
Db 584 LYKKNRYALKSGGSLNPNPENGQTNNDWILMGSTQEEAKKNMKNRRIGDFGG 643

Qy 61 PFDEENGKNGALNPNFNGKSAQNRFLLTGGANLNG-----GN-----GRP 102
Db 644 PFGEENGKNGALNPNFNGKSAQNRFLLTGGTNLNGKISVTQGNVLLSGRP 695

RESULT 11
Q9JVB9 PRELIMINARY; PRT; 1773 AA.
AC Q9JVB9;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TRENBLrel. 15, Last annotation update)
DE IGAL PROTEASE (EC 3.4.21.7).
GN IGA OR NMA0905.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Fellwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;

```

RT "Complete DNA sequence of a serogroup A strain of Neisseria
 RL meningitidis 22491."
 DR Nature 404:502-506(2000).
 KW EMBL: AL162754; CAB84182.1; -
 SQ Protease; Hydrolase.
 SEQUENCE 1773 AA; 196350 MW; CAC19E7313D76CE1 CRC64;

Query Match 80.5%; Score 458; DB 2; Length 1773;
 Best Local Similarity 77.7%; Pred. No. 1.2e-38;
 Matches 87; Conservative 5; Mismatches 10; Indels 10; Gaps 2;
 Qy 1 LYXNRYALKSGRLNAPNGVNAENDWPMGYTQEQEARKNAMNNKNNRIGDFGG 60
 Db LYXNRYALKSGVSNAPNGVNAENDWPMGYTQEQEARKNAMNNKNNRIGDFGG 60
 Qy 61 FFEENGKGHGALNLFNGKSAQNRELLTGGANLNG-----GN-----GRP 102
 Db 633 FFEENGKGHGALNLFNGKSAQNRELLTGGTNLNGKISVTQGNVLLSGRP 684

RESULT 12
 Q9K0B4
 ID Q9K0B4 PRELIMINARY; PRT; 1815 AA.
 AC Q9K0B4;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE ICA-SPECIFIC SERINE ENDOPEPTIDASE.
 GN NMB0700.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 ON NCBI_TaxID=491;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Vamathevan J.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58."
 RL Science 287:1809-1815(2000).
 DR EMBL: AE002424; AAF41117.1; -
 DR TIGR; NMB0700; -
 SQ SEQUENCE 1815 AA; 201077 MW; 2259D4D71762C57F CRC64;

Query Match 79.8%; Score 454; DB 2; Length 1815;
 Best Local Similarity 76.8%; Pred. No. 3.3e-38;
 Matches 86; Conservative 6; Mismatches 10; Indels 10; Gaps 2;
 Qy 1 LYXNRYALKSGRLNAPNGVNAENDWPMGYTQEQEARKNAMNNKNNRIGDFGG 60
 Db LYXNRYALKSGVSNAPNGVNAENDWPMGYTQEQEARKNAMNNKNNRIGDFGG 60
 Qy 61 FFEENGKGHGALNLFNGKSAQNRELLTGGANLNG-----GN-----GRP 102
 Db 633 FFEENGKGHGALNLFNGKSAQNRELLTGGTNLNGKISVTQGNVLLSGRP 684

RESULT 13
 Q9JXK3
 ID Q9JXK3 PRELIMINARY; PRT; 1431 AA.
 AC Q9JXK3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DE ICA-SPECIFIC SERINE ENDOPEPTIDASE.
 GN NMB0700.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 ON NCBI_TaxID=491;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Vamathevan J.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58."
 RL Science 287:1809-1815(2000).
 DR EMBL: AE002424; AAF41117.1; -
 DR TIGR; NMB0700; -
 SQ SEQUENCE 1815 AA; 201077 MW; 2259D4D71762C57F CRC64;

DE SERINE-TYPE PEPTIDASE.
 GN NMB1998.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 ON NCBI_TaxID=491;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Vamathevan J.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58."
 RL Science 287:1809-1815(2000).
 DR EMBL: AE002549; AAF42325.1; -
 DR TIGR; NMB1998; -
 SQ SEQUENCE 1431 AA; 157645 MW; 6801F0E209C24EFC CRC64;

Query Match 23.6%; Score 134; DB 2; Length 1431;
 Best Local Similarity 28.8%; Pred. No. 2.1e-05;
 Matches 42; Conservative 18; Mismatches 40; Indels 46; Gaps 7;
 Qy 3 YKNYR--YALKSGRLNAPNGVNAENDWPMGYTQEQEARKNAMNNKNNRIGDFG 59
 Db 555 HRNRRTDYFILKPGGNPREFFPLN-MKNSTSWQFIGNRQQAARQAENARPDLTFF 613
 Qy 60 GFDE--ENGKG-----HNGALNLFNGKSAQNRF 87
 Db 614 GYGENAQTGRAPSYSKTNEAEKTRHTANAAVYGRPEYRYGALNLHYRPRKRTDSTL 673
 Qy 88 LLTGGANL-----GGN-----GRP 103
 Db 674 LLNGMNLNGEVLEGGNMIVSGRPV 699

RESULT 14
 Q9X7H1
 ID Q9X7H1 PRELIMINARY; PRT; 1457 AA.
 AC Q9X7H1;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE APP PROTEIN.
 GN APP.
 OS Neisseria meningitidis serogroup B.
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 ON NCBI_TaxID=491;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Vamathevan J.,
 RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58."
 RL Science 287:1809-1815(2000).
 DR EMBL: AE002424; AAF41117.1; -
 DR TIGR; NMB0700; -
 SQ SEQUENCE 1457 AA; 159956 MW; CE62390B6C04B781 CRC64;

Query Match 15.9%; Score 90.5; DB 2; Length 1457;
 Best Local Similarity 25.6%; Pred. No. 0.64;
 Matches 32; Conservative 20; Mismatches 38; Indels 35; Gaps 7;
 Qy 8 YALKSGRLNAPNGVNA-----ENND--WVFNGYTQEE-----ARKNAMNNKNN 52

```
Db 527 YFGFR-GGRDL-NGHSLSFHRIQNTDEGAMIVNHNQDKESTVTITGNKDIATTGNN 581
QY 53 RRIGD-----FGGFFDEENGKGNALNLFNGKSAQNRFLLTGKANLNGG----- 98
Db 582 NSLDSKKEITAYNGWGEKDTTNGRLNLVYOPAAEDRTLLSSGGTNLNGNITQTNGKLF 641
QY 99 -NGRP 102
Db 642 FSGRP 646

RESULT 15
Q9JXL6
ID Q9JXL6 PRELIMINARY; PRT: 1457 AA.
AC Q9JXL6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ADHESION AND PENETRATION PROTEIN.
GN NMB1985.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Cittone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT *Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58."
RL Science 287:1809-1815(2000).
DR EMBL: AF002547; AAF42312.1; -.
DR TIGR: NMB1985; -.
SQ SEQUENCE 1457 AA; 159965 MW; B759529CFD4BD0AF CRC64;
```

```
Query Match 15.9%; Score 90.5; DB 2; Length 1457;
Best Local Similarity 25.6%; Pred. No. 0.64;
Matches 32; Conservative 20; Mismatches 38; Indels 35; Gaps 7;

QY 8 YALKSGGRLNAPMPENGVA-----ENND--WVFNGYTOEE-----ARKNAMNNKNN 52
Db 527 YFGFR-GGRDL-NGHSLSFHRIQNTDEGAMIVNHNQDKESTVTITGNKDIATTGNN 581
QY 53 RRIGD-----FGGFFDEENGKGNALNLFNGKSAQNRFLLTGKANLNGG----- 98
Db 582 NSLDSKKEITAYNGWGEKDTTNGRLNLVYOPAAEDRTLLSSGGTNLNGNITQTNGKLF 641
QY 99 -NGRP 102
Db 642 FSGRP 646
```

Search completed: May 1, 2001, 15:04:55
Job time: 321 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 14:58:50 ; Search time 66.74 Seconds
(without alignments)
89.077 Million cell updates/sec

Title: US-09-142-970-5

Perfect score: 565

Sequence: 1 LYKRYRYALKSGRLNAP.....NRFLTGGANLNGKISVTQG 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	565	100.0	104	19 W61606	Neisseria IgA1 pro
2	524	92.7	104	19 W61605	Neisseria IgA1 pro
3	514	91.0	104	19 W61603	Neisseria IgA1 pro
4	501	88.7	104	19 W61604	Neisseria IgA1 pro
5	494	87.4	105	19 W65656	105-mer peptide us
6	494	87.4	105	19 W61602	Neisseria IgA1 pro
7	298.5	52.8	1541	11 R07304	IgA1 protease. Ha
8	132	23.4	1431	21 Y75564	Neisseria meningit
9	115.5	20.4	1468	20 Y38825	Neisseria gonorrhoe
10	101	17.9	1394	17 R92768	Adhesion and penet
11	99.5	17.6	1454	21 Y56621	Neisseria meningit

12	99.5	17.6	1457	20 Y38823	Neisseria meningit
13	99.5	17.6	1457	21 B58592	N. meningitidis am
14	99.5	17.6	1457	21 B25662	N. meningitidis am
15	99.5	17.6	1457	21 Y56622	Neisseria meningit
16	93.5	16.5	1449	20 Y38824	Neisseria meningit
17	75	13.3	461	16 R75506	Nicotiana glauca
18	72.5	12.8	597	12 R12384	Bovine pancreatic
19	72	12.7	1221	21 B01825	Haemophilus influe
20	72	12.7	1227	21 B01824	Haemophilus influe
21	70	12.4	1577	17 R91047	Alpha-D-glucosyltr
22	69.5	12.3	263	21 G42442	Arabidopsis thalia
23	69.5	12.3	263	21 G57383	Arabidopsis thalia
24	69.5	12.3	263	21 G59074	Arabidopsis thalia
25	69.5	12.3	287	21 G42441	Arabidopsis thalia
26	69.5	12.3	288	21 G57382	Arabidopsis thalia
27	68.5	12.1	631	16 R77894	Bacterial transfer
28	68.5	12.1	1247	21 B18215	Plasmodium falcipa
29	68	12.0	834	20 Y34542	Porphorymonas ging
30	68	12.0	907	20 Y34408	Porphorymonas ging
31	68	12.0	1222	21 B01830	H. influenzae stra
32	68	12.0	1228	21 B01828	Haemophilus influe
33	67.5	11.9	631	18 W08968	Amino acid sequenc
34	67.5	11.9	631	19 W54126	H. influenzae stra
35	67.5	11.9	631	21 Y51782	H. influenzae non-
36	67.5	11.9	631	21 Y80364	H. influenzae stra
37	67.5	11.9	764	21 B18282	Plasmodium falcipa
38	67.5	11.9	807	21 B18311	Plasmodium falcipa
39	67	11.9	537	7 P60452	Sequence of the As
40	67	11.9	722	20 Y43379	M. catarrhalis str
41	67	11.9	1817	21 B18255	Plasmodium falcipa
42	67	11.9	2893	19 W98828	H. pylori GHP0 148
43	67	11.9	2893	19 W71556	Helicobacter polyp
44	66.5	11.8	631	16 R77896	Bacterial transfer
45	66.5	11.8	631	18 W08970	Amino acid sequenc

ALIGNMENTS

RESULT 1

W61606
ID W61606 standard; peptide; 104 AA.

XX AC W61606;

XX AC W61606;

DT 27-OCT-1998 (first entry)

XX DE Neisseria IgA1 protease fragment 5.

XX DE Immunoglobulin protease; carrier; paediatric; vaccine;

KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.

XX OS Neisseria sp.

XX OS Neisseria sp.

PN WO9831791-A1.

XX PD 23-JUL-1998.

XX PF 20-JAN-1998; 98WO-EP00294.

XX PR 21-JAN-1997; 97EP-0100883.

XX PR 21-JAN-1997; 97EP-0100883.

XX PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX PI Achtmann M, Moreau M;

XX PI 105-mer peptide us

XX DR Neisseria IgA1 pro

XX DR IgA1 protease. Ha

XX PT New peptide from Neisseria immunoglobulin protease - useful as

PT immunogenic carrier, e.g. particularly for polysaccharide(s),

PT forming conjugates used in vaccines against Neisseria and

PT Haemophilus

XX PS Claim 2; Fig 4; 32pp; English.

XX CC The Neisseria immunoglobulin protease fragments W61602-W61606 are used

CC as carriers for a conjugate, particularly in combination with a

CC polysaccharide. They can be used in paediatric or other vaccines,

CC particularly for prevention of epidemic bacterial infections, especially

CC those caused by Neisseria or Haemophilus. The protease fragment is a

CC highly immunogenic carrier that elicits a T-cell response, resulting in a

CC long-lasting memory and high antibody titre, and possibly making possible

CC vaccination without adjuvant.

XX SQ Sequence 104 AA;

Query Match 100.0%; Score 565; DB 19; Length 104;

Best Local Similarity 100.0%; Pred. No. 8.1e-59;

Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYYKNRYALKSGGRLNAPMPENGVAENNDFWFGYTQEEARKNAMHKNRRIGDFGG 60

Db 1 lyyknyryalksggrlnapmpengvaenndfwfgytqeearknamhknrrigdfgg 60

QY 61 FFDEENGKGHNGALNLFNGKSAQNRFLLTGGANLNGKISVTQG 104

Db 61 ffdeengkgngalnlnfngksaqrnlftgganlngkivtgg 104

RESULT 2

W61605

ID W61605 standard; peptide; 104 AA.

XX W61605;

XX 27-OCT-1998 (first entry)

XX Neisseria IgA1 protease fragment 4.

XX Immunoglobulin protease; carrier; paediatric; vaccine;

KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.

XX Neisseria sp.

XX WO9831791-Al.

XX 23-JUL-1998.

XX 20-JAN-1998; 98WO-EP00294.

XX 21-JAN-1997; 97EP-0100883.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX Achtmann M, Moreau M;

XX WPI; 1998-414092/35.

XX New peptide from Neisseria immunoglobulin protease - useful as

PT immunogenic carrier, e.g. particularly for polysaccharide(s),

PT forming conjugates used in vaccines against Neisseria and

PT Haemophilus

XX Claim 2; Fig 1/4; 32pp; English.

XX The Neisseria immunoglobulin protease fragments W61602-W61606 are used

CC as carriers for a conjugate, particularly in combination with a

CC polysaccharide. They can be used in paediatric or other vaccines,

CC particularly for prevention of epidemic bacterial infections, especially

CC those caused by Neisseria or Haemophilus. The protease fragment is a

CC highly immunogenic carrier that elicits a T-cell response, resulting in a

CC long-lasting memory and high antibody titre, and possibly making possible

CC vaccination without adjuvant.

XX SQ Sequence 104 AA;

Query Match 92.7%; Score 524; DB 19; Length 104;

Best Local Similarity 97.9%; Pred. No. 5.1e-54;

Matches 95; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYYKNRYALKSGGRLNAPMPENGVAENNDFWFGYTQEEARKNAMHKNRRIGDFGG 60

Db 1 lyyknyryalksggrlnapmpengvaenndfwfgytqeearknamhknrrigdfgg 60

QY 61 FFDEENGKGHNGALNLFNGKSAQNRFLLTGGANLNG 97

Db 61 ffdeengkgngalnlnfngksaqrnlftgganlng 97

RESULT 3

W61603

ID W61603 standard; peptide; 104 AA.

XX W61603;

XX 27-OCT-1998 (first entry)

XX Neisseria IgA1 protease fragment 2.

XX Immunoglobulin protease; carrier; paediatric; vaccine;

KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.

XX Neisseria sp.

XX WO9831791-Al.

XX 23-JUL-1998.

XX 20-JAN-1998; 98WO-EP00294.

XX 21-JAN-1997; 97EP-0100883.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX Achtmann M, Moreau M;

XX WPI; 1998-414092/35.

XX New peptide from Neisseria immunoglobulin protease - useful as

PT immunogenic carrier, e.g. particularly for polysaccharide(s),

PT forming conjugates used in vaccines against Neisseria and

PT Haemophilus

XX Claim 2; Fig 2; 32pp; English.

XX The Neisseria immunoglobulin protease fragments W61602-W61606 are used

CC as carriers for a conjugate, particularly in combination with a

CC polysaccharide. They can be used in paediatric or other vaccines,

CC particularly for prevention of epidemic bacterial infections, especially

CC those caused by Neisseria or Haemophilus. The protease fragment is a

CC highly immunogenic carrier that elicits a T-cell response, resulting in a

CC long-lasting memory and high antibody titre, and possibly making possible

CC vaccination without adjuvant.

XX SQ Sequence 104 AA;

Query Match 91.0%; Score 514; DB 19; Length 104;

Best Local Similarity 89.4%; Pred. No. 7.6e-53;

Matches 93; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 LYYKNRYALKSGGRLNAPMPENGVAENNDFWFGYTQEEARKNAMHKNRRIGDFGG 60

Db 1 lyyknyryalksggrlnapmpengvaenndfwfgytqeearknamhknrrigdfgg 60

XX IgAl; vaccine; meningitis; gonorrhoea; allergies.
 KW Haemophilus influenzae.
 XX OS
 XX W09011367-A.
 XX PPN
 XX PD 04-OCT-1990.
 XX 16-MAR-1990; 90WO-DK00073.
 XX PF
 XX 17-MAR-1989; 89DK-0001308.
 XX PPR
 XX PA (KILL/) KILIAN M.
 XX PI
 XX PI Kilian M, Poulsen K;
 XX PPI
 XX DR WPI; 1990-320267/42.
 XX DR N-PSDB; Q06164.
 XX PPT
 XX PPT Immunoglobulin A1 protease prodn. - by cloning from
 XX PPT microorganisms for immunisation against immunoglobulin A1
 XX PPT protease producing bacteria
 XX PS
 XX PS Disclosure; fig 3; 44pp; English.
 XX CCC
 XX CCC This immunoglobulin (Ig)A1 protease is produced by recombinant DNA
 XX CCC methods. It is useful in a vaccine for e.g. meningococcal meningitis,
 XX CCC gonorrhoea or allergic diseases. It specifically cleaves the heavy
 XX CCC chain of human IgA1 in the hinge region.
 XX CC
 XX CC Sequence 1541 AA;
 XX SQ

Query Match 52.8%; Score 298.5; DB 11; Length 1541;
 Best-Local Similarity 51.9%; Pred. No. 4.4e-26;
 Matches 54; Conservative 18; Mismatches 31; Indels 1; Gaps

QY 1 L YKVKRYVALKSGGRLNAPMPENGVAENNDWIFMGYTOEARKNAMHKNNRRIGDFGG 60
 DB 595 lnlenyyalrkagstselpkngsgesnewlymgktsdeakrnvmhinnermgng 654

QY 61 FDEENGKNGHNGALNINFGKSAQNRFLLTGGANLNGKISVTQG 104
 DB 655 yfgeeeek-ngnlnvtrfkqseqnrfliatggtnlmgdlvtvek 697

RESULT 8
 ID Y75564 standard; Protein; 1431 AA.
 XX AC
 XX AC Y75564;
 XX AC
 XX 21-MAR-2000 (first entry)
 XX DE
 XX DE Neisseria meningitidis ORF 759 protein sequence SEQ ID NO:2602.
 XX KW
 XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
 XX KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
 XX KW antibacterial; gene therapy.
 XX OS
 XX OS Neisseria meningitidis.
 XX PN
 XX PN W09957280-A2.
 XX PD
 XX PD 11-NOV-1999.
 XX PF
 XX PF 30-APR-1999; 99WO-US09346.
 XX PR
 XX PR 01-MAY-1998; 98US-0083758.
 XX PR 31-JUL-1998; 98US-0094869.
 XX PR 02-SEP-1998; 98US-0098994.
 XX PR 02-SEP-1998; 98US-0099062.

FT Active-site 243 /note= "Active site serine residue"
 FT Region 698 /note= "Conserved cysteine residue"
 FT Region 709 /note= "Conserved cysteine residue"
 FT Cleavage-site 960 /note= "Site for cleavage of secreted protease from outer membrane protein"
 FT Protein 960..1394
 FT /note= "45-kDa outer membrane protein fragment"
 XX W09605858-A1.
 XX 29-FEB-1996.
 XX 16-AUG-1995; 95WO-US10661.
 XX 25-AUG-1994; 94US-0296791.
 XX (STRD) UNIV LELAND STANFORD JUNIOR.
 XX (UNIW) UNIV WASHINGTON.
 XX Falkow S, St Geme JW;
 XX WPI; 1996-151147/15.
 XX N-PSDB; T17215.
 XX Haemophilus adhesion and penetration protein and corresponding DNA
 FT - used to produce vaccines against H. influenzae infection
 XX Claim 2; Fig 6; 105pp; English.
 XX The sequence represents a Haemophilus influenzae adhesion and
 CC penetration protein, encoded by the nap gene. The protein is
 CC first synthesised as a preprotein with 3 functional domains (the
 CC N-terminal signal peptide, the protease and a C-terminal helper
 CC domain), which is transported to the periplasm, followed by
 CC insertion of the C-terminal beta-domain into the outer membrane,
 CC possibly forming a pore, and export of the N-terminal fragment
 CC through the outer membrane, followed by autoproteolytic cleavage
 CC and secretion of the mature protease, leaving an outer membrane
 CC protein fragment. An active site serine residue characteristic of
 CC serine proteases is present, along with 2 cysteine residues which
 CC may be important in forming the catalytic domain. A gene encoding
 CC the protein may be inserted in a vector and expressed in
 CC recombinant host cells, for use as a recombinant vaccine. The
 CC protein may also be used in diagnostic monoclonal antibody production.
 XX Sequence 1394 AA;

Query Match 17.9%; Score 101; DB 17; Length 1394;
 Best Local Similarity 36.7%; Pred. No. 0.0054;
 Matches 22; Conservative 11; Mismatches 27; Indels 0; Gaps 0;
 Qy 45 NAMHKNNRRIGDGGFFDEENGKNGHGNALNPNFGSAQNRRLLTGAGNKGISVYTOG 104
 I : : I : : I : : I : : I : : I : : I : : I : : I : : I : : I : :
 Db 551 nninkldyrkeiayngwfgedtknkhgrlnliykpttedrlllsggtnlkgdtkg 610
 RESULT 11
 Y56621
 ID Y56621 standard; Protein; 1454 AA.
 XX AC Y56621;
 XX 21-FEB-2000 (first entry)
 XX Neisseria meningitidis strain ATCC13090 BASB006 protein sequence.
 XX Neisseria meningitidis; BASB006; diagnosis; bacterial; infection;
 KW vaccine# antibiotic; upper respiratory tract infection; meningitis;

KW invasive bacterial disease; bacteraemia; screening; antibacterial.
 XX Neisseria meningitidis.
 XX W09955873-A2.
 XX 04-NOV-1999.
 XX 20-APR-1999; 99WO-EP02766.
 XX 24-APR-1998; 98GB-0008866.
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX Thonnard J;
 XX WPI; 2000-052810/04.
 XX N-PSDB; Z38917.
 XX Novel polynucleotides and polypeptides from Neisseria meningitis used
 FT to prepare vaccines against bacterial infections
 XX Claim 3; Page 86-91; 103pp; English.
 XX The present sequence is BASB006 isolated from Neisseria meningitidis
 CC strain ATCC13090. BASB006 polynucleotides and polypeptides may be
 CC employed as research reagents and material for the discovery of
 CC treatments and diagnostics for diseases, particularly human diseases.
 CC They can be used for diagnosis of disease, staging of disease, or
 CC determining response of an infectious organism to drugs. The
 CC polynucleotides may be used as a source for hybridisation probes, and for
 CC screening of genetic mutations, serotype, organism or strain
 CC identification, identification of mutation in BASB013 sequences, and as
 CC components of arrays which are useful for diagnostic and prognostic
 CC purposes. The polypeptides can be used to produce antibodies. The
 CC polypeptides can also be used in vaccine formulations, and to identify
 CC agonists and antagonists. The polypeptides, antibodies, agonists and
 CC antagonists (which are bacteriostatic) are used for the treatment and
 CC prevention of diseases such as upper respiratory tract infection,
 CC invasive bacterial diseases such as bacteraemia and meningitis, and for
 CC the development and screening of antibacterial drugs. They are also used
 CC in the prevention of adhesion of bacteria to eukaryotic matrix proteins
 CC on in-dwelling devices, or to extracellular proteins on wounds, and to
 CC prevent tissue damage and/or block the normal progression of pathogenesis
 CC in infections initiated other than by the implantation of in-dwelling
 CC devices or by other surgical techniques.
 XX Sequence 1454 AA;

Query Match 17.6%; Score 99.5; DB 21; Length 1454;
 Best Local Similarity 27.4%; Pred. No. 0.0085;
 Matches 32; Conservative 18; Mismatches 42; Indels 25; Gaps 4;
 Qy 8 YYALKSGGRNLNAPENGVA-----ENNDFWFMGYTQDEARKNAMHKNRRRTGDFG-- 59
 I : : I I I : : I : : I : : I : : I : : I : : I : : I : : I : :
 Db 527 yfgfr-ggridl-----nghsisfrhqtdegamivnhdqkdestvtitgnkdattgnn 581
 Qy 60 -----GFFDEENGKNGHGNALNPNFGSAQNRRLLTGAGNKGISVYTOG 104
 I : : I : : I : : I : : I : : I : : I : : I : : I : : I : : I : :
 Db 582 nlsdskelayngwfgedtknkhgrlnliykpttedrlllsggtnlkgdtkg 638
 RESULT 12
 Y38823
 ID Y38823 standard; Protein; 1457 AA.
 XX AC Y38823;
 XX 08-OCT-1999 (first entry)
 XX Neisseria meningitidis antigen encoded by ORF1.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 14:58:48 ; Search time 66.74 Seconds
(without alignments)
89.077 Million cell updates/sec

Title: US-09-142-970-3

Perfect score: 563

Sequence: 1 LYKKNRYALKSGSVNAP.....NRFLTGTNLNGKISVTQG 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	563	100.0	104	19 W61604	Neisseria IgA1 pro
2	546	97.0	104	19 W61603	Neisseria IgA1 pro
3	537	95.4	105	19 W61606	105-mer peptide us
4	537	95.4	105	19 W61602	Neisseria IgA1 pro
5	501	89.0	104	19 W61606	Neisseria IgA1 pro
6	462	82.1	104	19 W61605	Neisseria IgA1 pro
7	370.5	56.9	1541	11 R07304	IgA1 protease. Ha
8	132	23.4	1431	21 Y75564	Neisseria meningit
9	127.5	22.6	1468	20 Y38825	Neisseria gonorrhoe
10	114	20.2	1394	17 R92768	Adhesion and penet
11	111.5	19.8	1449	20 Y38824	Neisseria meningit

12	109	19.4	1454	21 Y56621	Neisseria meningit
13	109	19.4	1457	20 Y38823	Neisseria meningit
14	109	19.4	1457	21 Y58592	N. meningitidis am
15	109	19.4	1457	21 B25862	N. meningitidis am
16	109	19.4	1457	21 Y56622	Neisseria meningit
17	78	13.9	461	16 R75506	Nicotiana glauca
18	78	13.9	1577	17 R91047	Alpha-D-glucosyltr
19	76.5	13.6	280	21 B63197	Gene 21 human secr
20	76.5	13.6	447	17 W01622	Oyster pearl prote
21	76.5	13.6	954	21 B48550	Ruminococcus flave
22	73.5	13.1	806	15 R55892	hnRNP U protein.
23	72	12.8	764	21 B18282	Plasmodium falcipa
24	71	12.6	394	19 W41729	Actinobacillus ple
25	70.5	12.5	631	16 R77896	Bacterial transfer
26	70.5	12.5	631	18 W08970	Amino acid sequenc
27	70.5	12.5	631	19 W54128	H. influenzae stra
28	70.5	12.5	631	21 Y51784	H. influenzae non-
29	69	12.3	275	21 Y80366	H. influenzae stra
30	69	12.3	288	21 G57194	Arabidopsis thalia
31	69	12.3	311	21 G57192	Arabidopsis thalia
32	69	12.3	537	7 P60452	Sequence of the As
33	68.5	12.2	1073	21 B01837	Haemophilus influe
34	68.5	12.2	1079	21 B01836	Haemophilus influe
35	68	12.1	487	21 Y83172	Cell wall protein
36	68	12.1	487	21 Y70121	Staph. epidermidis
37	67.5	12.0	437	16 R75507	Nicotiana plumbagi
38	67.5	12.0	611	17 R88653	N. meningitidis M97
39	67.5	12.0	611	17 R88653	N. meningitidis M97
40	67.5	12.0	671	17 R85290	Streptococcus faec
41	67.5	12.0	705	17 R88645	Neisseria meningit
42	67	11.9	546	17 W05406	Mouse cortactin pr
43	67	11.9	1253	21 B19772	Human dentin sialo
44	66.5	11.8	776	20 Y34575	Porphyromonas ging
45	66.5	11.8	814	20 Y34432	Porphyromonas ging

ALIGNMENTS

RESULT 1

W61604
ID W61604 standard; peptide; 104 AA.
AC W61604;
XX
XX
DF 27-OCT-1998 (first entry)
XX
DE Neisseria IgA1 protease fragment 3.
XX
KW Immunoglobulin protease; carrier; paediatric; vaccine;
KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.
XX
OS Neisseria sp.
XX
PN W09831791-A1.
XX
PD 23-JUL-1998.
XX
PF 20-JAN-1998; 98WO-EP00294.
XX
PR 21-JAN-1997; 97EP-0100883.
XX
PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
XX
PI Achtmann M, Moreau M;
XX
DR WPI; 1998-414092/35.
XX
XX

New peptide from Neisseria immunoglobulin protease - useful as
PT immunogenic carrier, e.g. particularly for polysaccharide(s).
PT forming conjugates used in vaccines against Neisseria and
PT Haemophilus

XX PS Claim 2; Fig 1; 32pp; English.

XX CC The Neisseria immunoglobulin protease fragments W61602-W61606 are used as carriers for a conjugate, particularly in combination with a polysaccharide. They can be used in paediatric or other vaccines, particularly for prevention of epidemic bacterial infections, especially those caused by Neisseria or Haemophilus. The protease fragment is a highly immunogenic carrier that elicits a T-cell response, resulting in a long-lasting memory and high antibody titre, and possibly making possible vaccination without adjuvant.

XX SQ Sequence 104 AA;

Query Match 100.0%; Score 563; DB 19; Length 104;
Best Local Similarity 100.0%; Pred. No. 7.5e-57;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYYKNRYALKSGSVNAPMPENGQTENNDFVMGYKOEAAQKNAMHNKRNORISGFSG 60
|||||
Db 1 Lyyknyryalksgsvnapmpengqtenndwvfmgykqeeaknamhknqrisgfsq 60

QY 61 PFGEENGKHNGALNLFNGKSAQNRFLLTGTTNNGKISVTQG 104
|||||
Db 61 ffeengkhngalnlnfngksaqrflitggtngklsvtqg 104

RESULT 2

W61603

ID W61603 standard; peptide; 104 AA.

XX AC W61603;

XX DT 27-OCT-1998 (first entry)

XX DE Neisseria IgA1 protease fragment 2.

XX KW Immunoglobulin protease; carrier; paediatric; vaccine;

XX KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.

XX OS Neisseria sp.

XX PN WO9831791-A1.

XX PD 23-JUL-1998.

XX PF 20-JAN-1998; 98WO-EP00294.

XX PR 21-JAN-1997; 97EP-0100883.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

XX PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX PI Achtmann M, Moreau M;

XX DR WPI; 1998-414092/35.

XX PT New peptide from Neisseria immunoglobulin protease - useful as immunogenic carrier, e.g. particularly for polysaccharide(s), forming conjugates used in vaccines against Neisseria and Haemophilus

XX PS Claim 2; Fig 2; 32pp; English.

XX CC The Neisseria immunoglobulin protease fragments W61602-W61606 are used as carriers for a conjugate, particularly in combination with a polysaccharide. They can be used in paediatric or other vaccines, particularly for prevention of epidemic bacterial infections, especially those caused by Neisseria or Haemophilus. The protease fragment is a highly immunogenic carrier that elicits a T-cell response, resulting in a long-lasting memory and high antibody titre, and possibly making possible vaccination without adjuvant.

XX SQ Sequence 104 AA;

Query Match 97.0%; Score 546; DB 19; Length 104;
Best Local Similarity 97.1%; Pred. No. 6.5e-55;
Matches 101; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LYYKNRYALKSGSVNAPMPENGQTENNDFVMGYKOEAAQKNAMHNKRNORISGFSG 60
|||||
Db 1 Lyyknyryalksgsvnapmpengqtenndwvfmgykqeeaknamhknqrisgfsq 60

QY 61 PFGEENGKHNGALNLFNGKSAQNRFLLTGTTNNGKISVTQG 104
|||||
Db 61 ffeengkhngalnlnfngksaqrflitggtngklsvtqg 104

RESULT 3

W65656

ID W65656 standard; peptide; 105 AA.

XX AC W65656;

XX DT 15-OCT-1998 (first entry)

XX DE 105-mer peptide used in polysaccharide-peptide conjugate.

XX KW Polysaccharide-peptide conjugate; vaccine; linker moiety; adjuvant;

XX KW immune response.

XX OS Synthetic.

XX FH Key Location/Qualifiers

XX FT Modified-site 1 /note= "N-terminal acetyl"

XX PN WO9831393-A2.

XX PD 23-JUL-1998.

XX PF 21-JAN-1998; 98WO-EP00654.

XX PR 21-JAN-1997; 97EP-0100884.

XX PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX PI Mistretta N, Moreau M;

XX DR WPI; 1998-413820/35.

XX PT Polysaccharide-peptide conjugates, useful as, e.g. vaccines - comprise peptide moiety with at least six amino acid residues,

XX PT polysaccharide chain with at least four repeat units, and linker moiety

XX PS Example 1; Page 14; 28pp; English.

XX CC The invention relates to: (A) polysaccharide-peptide conjugate (in which the polysaccharide is immunogenic), comprising: (a) a peptide moiety which has at least 6 amino acid residues, at least 1 of which is a cysteine residue; (b) a polysaccharide chain comprising at least 4 repeat units, and (c) a linker moiety bound to the thiol group of the cysteine. The linker is also bound to: (i) native amino, hydroxyl or carboxyl groups of the polysaccharide chain; (ii) amino groups created by hydrolysis of native N-acyl groups of the polysaccharide chain, or (iii) functional groups introduced on the polysaccharide chain upon derivatisation with a spacer moiety bound to native amino, hydroxyl or carboxyl groups of the polysaccharide chain, and (B) conjugating a peptide as in (Aa) to a polysaccharide chain comprising at least 4 repeat units, comprising: (a) coupling the peptide to a linker through the thiol group of the cysteine residue, and (b) coupling the linker to the polysaccharide chain through one of groups (i), (ii) or (iii) as described in (A). The conjugates are especially useful as vaccines to

CC elicit a protective long term immune response against a pathogenic
 CC microorganism from which the immunogenic polysaccharide chain is derived.
 CC known polysaccharide-peptide conjugates are less immunogenic than their
 CC polysaccharide-protein counterparts and require adjuvantation. The new
 CC conjugates have good immunogenicity, largely because the peptides are of
 CC sufficient size. The present sequence represents a peptide used in
 CC a polysaccharide-peptide conjugate.

XX SQ Sequence 105 AA;

Query Match 95.4%; Score 537; DB 19; Length 105;
 Best Local Similarity 95.2%; Pred. No. 7e-54;
 Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LYYKNRYRYALKSGSVNAPMPENGQTEENNDFWPMGYKQEEAOKNAMHKNQRISGFSG 60
 |||||
 Db 2 lyyknyryyalksgsvnapmpengqtenndwilmgstqeeaknamhknqrsgisfg 61
 |||||

Qy 61 FFGENGKGHNGALNLFNGKSAQNRFLLTGTTNLNGKISVTQ 104
 |||||
 Db 62 ffeengkgghngalnlnfngksaqrflttgttnlنگkisvtqg 105
 |||||

RESULT 4

W61602 ID W61602 standard; peptide; 105 AA.

XX AC W61602;

XX DT 27-OCT-1998 (first entry)

XX DE Neisseria IgA1 protease fragment 1.

XX KW Immunoglobulin protease; carrier; paediatric; vaccine;
 KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.

XX OS Neisseria sp.

XX PN W09831791-A1.

XX PD 23-JUL-1998.

XX PF 20-JAN-1998; 98WO-EP00294.

XX PR 21-JAN-1997; 97EP-0100883.

XX PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX PI Achtmann M, Moreau M;

XX DR WPI; 1998-414092/35.

XX PT New peptide from Neisseria immunoglobulin protease - useful as
 PT immunogenic carrier, e.g. particularly for polysaccharide(s),
 PT forming conjugates used in vaccines against Neisseria and
 Haemophilus

XX PS Claim 6; Page 10; 32pp; English.

XX CC The Neisseria immunoglobulin protease fragments W61602-W61606 are used
 CC as carriers for a conjugate, particularly in combination with a
 CC polysaccharide. They can be used in paediatric or other vaccines,
 CC particularly for prevention of epidemic bacterial infections, especially
 CC those caused by Neisseria or Haemophilus. The protease fragment is a
 CC highly immunogenic carrier that elicits a T-cell response, resulting in a
 CC long-lasting memory and high antibody titre, and possibly making possible
 CC vaccination without adjuvant.

XX SQ Sequence 105 AA;

Query Match 95.4%; Score 537; DB 19; Length 105;
 Best Local Similarity 95.2%; Pred. No. 7e-54;
 Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LYYKNRYRYALKSGSVNAPMPENGQTEENNDFWPMGYKQEEAOKNAMHKNQRISGFSG 60
 |||||
 Db 2 lyyknyryyalksgsvnapmpengqtenndwilmgstqeeaknamhknqrsgisfg 61
 |||||

Qy 61 FFGENGKGHNGALNLFNGKSAQNRFLLTGTTNLNGKISVTQ 104
 |||||

Db 62 ffeengkgghngalnlnfngksaqrflttgttnlنگkisvtqg 105
 |||||

RESULT 5

W61606 ID W61606 standard; peptide; 104 AA.

XX AC W61606;

XX DT 27-OCT-1998 (first entry)

XX DE Neisseria IgA1 protease fragment 5.

XX KW Immunoglobulin protease; carrier; paediatric; vaccine;
 KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.

XX OS Neisseria sp.

XX PN W09831791-A1.

XX PD 23-JUL-1998.

XX PF 20-JAN-1998; 98WO-EP00294.

XX PR 21-JAN-1997; 97EP-0100883.

XX PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX PI Achtmann M, Moreau M;

XX DR WPI; 1998-414092/35.

XX PT New peptide from Neisseria immunoglobulin protease - useful as
 PT immunogenic carrier, e.g. particularly for polysaccharide(s),
 PT forming conjugates used in vaccines against Neisseria and
 Haemophilus

XX PS Claim 2; Fig 4; 32pp; English.

XX CC The Neisseria immunoglobulin protease fragments W61602-W61606 are used
 CC as carriers for a conjugate, particularly in combination with a
 CC polysaccharide. They can be used in paediatric or other vaccines,
 CC particularly for prevention of epidemic bacterial infections, especially
 CC those caused by Neisseria or Haemophilus. The protease fragment is a
 CC highly immunogenic carrier that elicits a T-cell response, resulting in a
 CC long-lasting memory and high antibody titre, and possibly making possible
 CC vaccination without adjuvant.

XX SQ Sequence 104 AA;

Query Match 89.0%; Score 501; DB 19; Length 104;
 Best Local Similarity 87.5%; Pred. No. 8.8e-50;
 Matches 91; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LYYKNRYRYALKSGSVNAPMPENGQTEENNDFWPMGYKQEEAOKNAMHKNQRISGFSG 60
 |||||
 Db 1 lyyknyryyalksggrinapmpengvaendwifmgvtqeeaknamhknrrigdfgg 60
 |||||

Qy 61 FFGENGKGHNGALNLFNGKSAQNRFLLTGTTNLNGKISVTQ 104
 |||||

Db 61 ffeengkgghngalnlnfngksaqrflttgttnlنگkisvtqg 104
 |||||

RESULT 6
W61605
ID W61605 standard; peptide; 104 AA.
XX W61605;
XX
DT 27-OCT-1998 (first entry)
XX
DE Neisseria IgA1 protease fragment 4.
XX
KW Immunoglobulin protease; carrier; paediatric; vaccine;
KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell
XX Neisseria sp.
XX OS
XX WO9831791-A1.
XX PN
XX PD 23-JUL-1998.
XX PF 20-JAN-1998; 98WO-EP00294.
XX PR 21-JAN-1997; 97EP-0100883.
XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA (INNR) PASTEUR MERIEUX SERUMS & VACCINS SA.
XX Achtmann M, Moreau M;
XX WPI; 1998-414092/35.
XX
XX New peptide from Neisseria immunoglobulin protease - useful as
PT immunogenic carrier, e.g. particularly for polysaccharide(s),
PT forming conjugates used in vaccines/against Neisseria and
PT Haemophilus
XX
XX Claim 2; Fig 1/4; 32pp; English.
XX
XX The Neisseria immunoglobulin protease fragments W61602-W61606 are used
CC as carriers for a conjugate, particularly in combination with a
CC polysaccharide. They can be used in paediatric or other vaccines,
CC particularly for prevention of epidemic bacterial infections, especially
CC those caused by Neisseria or Haemophilus. The protease fragment is a
CC highly immunogenic carrier that elicits a T-cell response, resulting in a
CC long-lasting memory and high antibody titre, and possibly making possible
CC vaccination without adjuvant.
XX
XX Sequence 104 AA;

Query Match 82.1%; Score 462; DB 19; Length 104;
Best Local Similarity 86.6%; Pred. No. 2.5e-45;
Matches 84; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
QY 1 LYKKNRYVYALKSGGSVNPMPENGOTENNDFVFMGYKOEAKNMMHNNRISGFSG 60
Db 1 lyyknyvyyalksggrlnapmpengvaennvfmgytqeaknmmknrrigdgfg 60
QY 61 FFGENGKCHGNALNLFNGKSAQNRFLLTGTGTLNG 97
Db 61 fideengkhgnalnlnfngksaqrflitgnganng 97

RESULT 7
R07304
ID R07304 standard; protein; 1541 AA.
XX
XX R07304;
XX
XX 31-JAN-1991 (first entry)
XX
XX IgA1 protease.

XX IgA1; vaccine; meningitis; gonorrhoea; allergies.
XX Haemophilus influenzae.
OS WO9011367-A.
PN
XX PD 04-OCT-1990.
XX PF 16-MAR-1990; 90WO-DK00073.
XX PR 17-MAR-1989; 89DK-0001308.
XX (KILL/) KILIAN M.
XX PI Kilian M, Poulsen K;
XX WPI; 1990-320267/42.
DR N-PSDB; Q06164.
XX Immunoglobulin A1 protease prodn. - by cloning from
PT microorganisms for immunisation against immunoglobulin A1
PT protease producing bacteria
XX Disclosure; fig 3; 44pp; English.
XX This immunoglobulin (Ig)A1 protease is produced by recombinant DNA
CC methods. It is useful in a vaccine for e.g. meningococcal meningitis,
CC gonorrhoea or allergic diseases. It specifically cleaves the heavy
CC chain of human IgA1 in the hinge region.
XX Sequence 1541 AA;

Query Match 56.9%; Score 320.5; DB 11; Length 1541;
Best Local Similarity 54.8%; Pred. No. 1.1e-27;
Matches 57; Conservative 21; Mismatches 25; Indels 1; Gaps 1;

QY 1 LYKKNRYVYALKSGGSVNPMPENGOTENNDFVFMGYKOEAKNMMHNNRISGFSG 60
Db 595 Inlenyvyalksgastrseipkngsgesnenwlyngktsdeakrnmhnmhnmngfng 654
QY 61 FFGENGKCHGNALNLFNGKSAQNRFLLTGTGTLNGKISVTOG 104
Db 655 yfgeeeegk-ngnlnvtfkgkseqnrflltgtgtnlmgdlitvekg 697

RESULT 8
Y75564
ID Y75564 standard; Protein; 1431 AA.
XX
XX Y75564;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 759 protein sequence SEQ ID NO:2602.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy.
XX
OS Neisseria meningitidis.
XX WO9957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
XX 01-MAY-1998; 98US-0083758.
XX 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.

```
PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
XX WPI; 2000-062150/05.
DR N-PSDB; Z54326.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics
XX
XX Claim 2; Page 1234; 1453pp; English.
XX
XX Z53015 to Z54536, Z54577 to Z54615, and Y74253 to Y75941 represent
CC novel Neisseria meningitis and N. gonorrhoeae polynucleotides and
CC polypeptides. Z54537 to Z54576 and Z54616 to Z54773 represent PCR
CC primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
XX Sequence 1431 AA;
SQ
Query Match 23.4%; Score 132; DB 21; Length 1431;
Best Local Similarity 26.3%; Pred. No. 3.1e-06;
Matches 36; Conservative 21; Mismatches 44; Indels 36; Gaps 5;
QY 3 YKNYR--YVALKSGGVNAPMPENGOTENNDFVPMGYKQEEAKNMMHKN-NQRISGFS 59
Db 555 hnrtdyflkpggnpreffpin-mkntswdfignrgaaevqaaenarpdlitfg 613
QY 60 GFPGF--ENGKG-----HNGALNLFNGKSAQNR 87
Db 614 gylgenaqtgkaapsysktnaektrhianaavgyrpeyrngalnlnhyrpkrdstl 673
QY 88 LITGGTNLNGKISVTQG 104
Db 674 lllgmgmlngevliegg 690
RESULT 9
Y38825
ID Y38825 standard; Protein; 1468 AA.
AC Y38825;
XX
XX 08-OCT-1999 (first entry)
DT
DE Neisseria gonorrhoeae antigenic protein encoded by ORF1.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicemia; gonorrhea.
XX
XX Neisseria gonorrhoeae.
OS
XX W09924578-A2.
PN
XX 20-MAY-1999.
PD
XX
```

```
PF 09-OCT-1998; 98WO-IB01665.
XX
XX 01-SEP-1998; 98GB-0019016.
PR 06-NOV-1997; 97GB-0023516.
PR 14-NOV-1997; 97GB-0024190.
PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX
XX (CHIR-) CHIRON SPA.
PA
XX Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
PI WPI; 1999-327407/27.
DR N-PSDB; Z12253.
XX
XX Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection
XX
XX Claim 4; Page 371; 524pp; English.
XX
XX Amino acid sequences Y38499-Y38944 represent Neisseria meningitidis
CC and N. gonorrhoeae antigenic proteins. They are encoded by open
CC reading frames (ORFs) Z1972-Z12358. The antigenic proteins,
CC their fragments, their nucleic acids and antibodies are used for
CC diagnosis, prevention (as vaccines) or treatment of Neisseria
CC infections, such as meningitis, septicemia and gonorrhea. Both
CC organisms are closely related. Fragments of the nucleic acids
CC are useful as hybridisation probes and antisense reagents.
XX
XX Sequence 1468 AA;
SQ
Query Match 22.6%; Score 127.5; DB 20; Length 1468;
Best Local Similarity 34.9%; Pred. No. 1e-05;
Matches 30; Conservative 16; Mismatches 39; Indels 1; Gaps 1;
QY 19 APMPENGOTENNDFVPMGYKQEEAKNMMHKN-NQRISGFSFGFENGKNGHGNALNLF 78
Db 554 amivnhnqkvestvitgkndittgtnnnldskkeia-yngwfgekdatktnglnlny 612
QY 79 NGKSAQNRFLLTGNTNLNGKISVTQG 104
Db 613 ppeeadrtllsgtngtnlgnitqng 638
RESULT 10
R92768
ID R92768 standard; Protein; 1394 AA.
XX
XX R92768;
XX
XX 01-JUN-1996 (first entry)
DT
DE Adhesion and penetration protein.
XX
XX Adhesion and penetration protein; hap gene; signal peptide;
KW protease; helper domain; pore; outer membrane protein; beta-domain;
KW secretion; active site; catalytic domain; recombinant vaccine;
KW monoclonal antibody; diagnostic; immunoassay.
XX
XX Haemophilus influenzae.
OS
XX
XX Key Location/Qualifiers
FH Peptide 1..25 /note= "signal peptide"
FT Peptide 27..33 /note= "Mature protein N-terminal peptide"
FT Protein 26..959 /note= "Secreted 110-kDa protease fragment"
FT Domain 241..248 /note= "Putative catalytic domain"
```



```

XX PF 20-APR-1999; 99WO-EP02766.
XX XX
XX PR 24-APR-1998; 98GB-0008866.
XX XX
XX PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX XX
XX PI Thonnard J;
XX XX
XX DR WPI; 2000-052810/04.
XX DR N-PSDB; Z38917.
XX XX
XX PT Novel polynucleotides and polypeptides from Neisseria meningitidis used
XX PT to prepare vaccines against bacterial infections
XX XX
XX PS Claim 3; Page 86-91; 103pp; English.
XX XX
XX CC The present sequence is BASB006 isolated from Neisseria meningitidis
XX CC strain ATCC13090. BASB006 polynucleotides and polypeptides may be
XX CC employed as research reagents and material for the discovery of
XX CC treatments and diagnostics for diseases, particularly human diseases.
XX CC They can be used for diagnosis of disease, staging of disease, or
XX CC determining response of an infectious organism to drugs. The
XX CC polynucleotides may be used as a source for hybridisation probes, and for
XX CC screening of genetic mutations, serotype, organism or strain
XX CC identification, identification of mutation in BASB013 sequences, and as
XX CC components of arrays which are useful for diagnostic and prognostic
XX CC purposes. The polypeptides can be used to produce antibodies. The
XX CC polypeptides can also be used in vaccine formulations, and to identify
XX CC agonists and antagonists. The polypeptides, antibodies, agonists and
XX CC antagonists (which are bacteriostatic) are used for the treatment and
XX CC prevention of diseases such as upper respiratory tract infection,
XX CC invasive bacterial diseases such as bacteraemia and meningitis, and for
XX CC the development and screening of antibacterial drugs. They are also used
XX CC on the prevention of adhesion of bacteria to eukaryotic matrix proteins
XX CC on in-dwelling devices, or to extracellular proteins on wounds, and to
XX CC prevent tissue damage and/or block the normal progression of pathogenesis
XX CC in infections initiated other than by the implantation of in-dwelling
XX CC devices or by other surgical techniques.
XX XX
XX SQ Sequence 1454 AA;

Query Match 19.4%; Score 109; DB 21; Length 1454;
Best Local Similarity 38.6%; Pred. No. 0.0013;
Matches 22; Conservative 13; Mismatches 22; Indels 0; Gaps 0;

QY 48 NHKNNORISGFSGFFGKNGHGNALNPNFGKSAQNRFLLTGGTNLNGKISVTQG 104
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 582 nsidskkelayngwfgkdtktnglnlvypqaadrtllsggtlnngnitqng 638

RESULT 13
Y38823
ID Y38823 standard; Protein; 1457 AA.
XX AC Y38823;
XX XX
XX DT 08-OCT-1999 (first entry)
XX XX
XX DE Neisseria meningitidis antigen encoded by ORF1.
XX XX
XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX KW treatment; Neisseria infection; meningitis; septicaemia; gonorrhea.
XX XX
XX OS Neisseria meningitidis.
XX XX
XX PN WO9924578-A2.
XX XX
XX PD 20-MAY-1999.
XX XX
XX PF 09-OCT-1998; 98WO-IB01665.
XX XX

```

```

PR 01-SEP-1998; 98GB-0019016.
PR 06-NOV-1997; 97GB-0023516.
PR 14-NOV-1997; 97GB-0024190.
PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX XX
XX PA (CHIR-) CHIRON SPA.
XX XX
XX PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX XX
XX DR WPI; 1999-327407/27.
XX DR N-PSDB; Z12251.
XX XX
XX PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
XX PT diagnosis, treatment and prevention of infection
XX PS Claim 4; Page 361; 524pp; English.
XX XX
XX CC Amino acid sequences Y38499-Y38944 represent Neisseria meningitidis
XX CC and N. gonorrhoeae antigenic proteins. They are encoded by open
XX CC reading frames (ORFs) Z1972-Z12358. The antigenic proteins,
XX CC their fragments, their nucleic acids and antibodies are used for
XX CC diagnosis, prevention (as vaccines) or treatment of Neisseria
XX CC infections, such as meningitis, septicaemia and gonorrhea. Both
XX CC organisms are closely related. Fragments of the nucleic acids
XX CC are useful as hybridisation probes and antisense reagents.
XX XX
XX SQ Sequence 1457 AA;

Query Match 19.4%; Score 109; DB 20; Length 1457;
Best Local Similarity 38.6%; Pred. No. 0.0013;
Matches 22; Conservative 13; Mismatches 22; Indels 0; Gaps 0;

QY 48 NHKNNORISGFSGFFGKNGHGNALNPNFGKSAQNRFLLTGGTNLNGKISVTQG 104
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 582 nsidskkelayngwfgkdtktnglnlvypqaadrtllsggtlnngnitqng 638

RESULT 14
B58592
ID B58592 standard; Protein; 1457 AA.
XX AC B58592;
XX XX
XX DT 13-MAR-2001 (first entry)
XX XX
XX DE N. meningitidis amino acid sequence orfl-1.pap SEQ ID NO:87.
XX XX
XX KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
XX KW diagnosis; antigen; detection; infection; gene therapy; antibacterial.
XX XX
XX OS Neisseria meningitidis.
XX XX
XX PN WO200066791-A1.
XX XX
XX PD 09-NOV-2000.
XX XX
XX PF 08-MAR-2000; 2000WO-0505928.
XX XX
XX PR 30-APR-1999; 99US-0132068.
XX PR 08-OCT-1999; 99WO-US23573.
XX PR 28-FEB-2000; 2000GB-0004695.
XX XX
XX PA (CHIR ) CHIRON CORP.
XX PA (GENO-) INST GENOMIC RES.
XX XX
XX PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;
XX PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
XX PI Frazer CM, Grandi G;
XX XX

```

```
DR WPI: 2000-647603/62.
XX N-PSDB; F21587.
PT Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections
XX
XX Example 1; Page 114; 692pp; English.
XX
CC The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in F21544 and F21607 to
CC F21613 represent fragments of the NMB genomic sequence, as the sequence
CC was too long to go in a record on its own it was split into 8 sequences
CC which overlap each other at the beginning and end of each sequence by
CC 49980 bp (i.e. the last 49980 bp of F21544 is repeated at the beginning of
CC F21607, the last 49980 bp of F21607 are repeated at the beginning of
CC F21608 and so on). F21545 to F21588 encode the Neisseria proteins given
CC in B58550 to B58593, and F21589 to F21606 represent PCR primers which
CC are used in the exemplification of the present invention. The NMB genome
CC and fragments from it have antibacterial activity, and can be used in
CC vaccines and gene therapy. Neisseria nucleic acids, proteins and/or
CC antibodies which binds to the proteins can be used in compositions for
CC treating or preventing infection due to Neisserial bacteria or as a
CC diagnostic reagent for detecting the presence of Neisserial bacteria or
CC of antibodies raised to Neisserial bacteria. Computers, computer memory,
CC computer storage medium or computer databases can be used in a search to
CC identify open reading frames (ORFs) or coding sequences within the NMB
CC genome. The DNA sequences provide further opportunities to find
CC antigenic or immunogenic proteins which are more effective in vaccines
XX Sequence 1457 AA;
SQ
Query Match 19.4%; Score 109; DB 21; Length 1457;
Best Local Similarity 38.6%; Pred. NO. 0.0013;
Matches 22; Conservative 13; Mismatches 22; Indels 0; Gaps 0;
QY 48 NHKNNQRISGFGFGEENGKNGHGNALNPNFGKSAQNRFLLTGGTNLNGKISVTQG 104
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 582 nslskskelayngwfgekdtktngnlvlyqpaaedrllllsggtlnngnitqng 638
RESULT 15
ID B25662 standard; Protein; 1457 AA.
XX
XX B25662;
XX
XX 04-DEC-2000 (first entry)
XX
XX N. meningitidis amino acid sequence orf1-1.pep SEQ ID NO:1047.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB.
XX
XX Neisseria meningitidis.
XX
XX WO200022430-A2.
XX
XX 20-APR-2000.
XX
XX 08-OCT-1999; 99WO-US23573.
XX
XX 09-OCT-1998; 98US-0103794.
XX
XX 30-APR-1999; 99US-0132068.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masiagnani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
XX WPI: 2000-318079/27.
DR
```

```
DR N-PSDB; A81302.
XX
XX Isolated nucleotide sequences of Neisseria meningitidis which can be
XX used in the diagnosis and treatment of N. meningitidis infection and
XX other Neisserial infections, for example, N.gonorrhoea
XX
XX Example 1; Page 113; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic
XX proteins from Neisseria genomic sequences. A81453 to A82414 represent
XX specifically claimed Neisseria meningitidis genomic DNA sequences; A81260
XX to A81303 and B25620 to B25663 represent Neisseria DNA sequences and
XX their corresponding proteins; A81254 to A81259 and A81304 to A81321
XX represent PCR primers used in the isolation of Neisseria meningitidis DNA
XX sequences; and A81322 to A81452 represent Neisseria meningitidis MenB
XX polynucleotide ORF sequences, which are all used in the exemplification
XX of the present invention. The nucleic acid sequences, protein sequences,
XX and antibodies against them, can be used in the manufacture of a
XX composition. The composition can be used as a medicament (or in the
XX manufacture of a medicament) for treating, preventing or diagnosing
XX infection due to Neisserial bacteria. For example, some of the identified
XX proteins could be components of vaccines against Meningococcus B; against
XX all serotypes; and/or against all pathogenic Neisseriae. Identification
XX of sequences from the bacterium will also facilitate production of
XX biological probes, particularly organism-specific probes. Attempts to
XX make efficacious Meningococcus B vaccines have failed mainly due to
XX antigen tolerance. Multivalent vaccines have also been tried but none
XX have successfully overcome antigenic variability. The provision of
XX further, complete sequences may provide an opportunity to identify
XX secreted or surface exposed proteins that may be presumed targets for the
XX immune system and which are not antigenically variable or at least more
XX conserved than other more variable regions.
XX Sequence 1457 AA;
SQ
Query Match 19.4%; Score 109; DB 21; Length 1457;
Best Local Similarity 38.6%; Pred. NO. 0.0013;
Matches 22; Conservative 13; Mismatches 22; Indels 0; Gaps 0;
QY 48 NHKNNQRISGFGFGEENGKNGHGNALNPNFGKSAQNRFLLTGGTNLNGKISVTQG 104
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 582 nslskskelayngwfgekdtktngnlvlyqpaaedrllllsggtlnngnitqng 638
Search completed: May 1, 2001, 14:58:49
Job time: 611 sec
```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 14:48:38 ; Search time 66.74 Seconds
(without alignments)
89.077 Million cell updates/sec

Title: US-09-142-970-1
Perfect score: 558
Sequence: 1 LYKNRYALKSGSVNAP.....NRLLTGTNLNGKISVTQG 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	558	100.0	105	19 W65656	105-mer peptide us
2	558	100.0	105	19 W61602	Neisseria IgA1 pro
3	540	96.8	104	19 W61603	Neisseria IgA1 pro
4	537	96.2	104	19 W61604	Neisseria IgA1 pro
5	494	88.5	104	19 W61606	Neisseria IgA1 pro
6	453	81.2	104	19 W61605	Neisseria IgA1 pro
7	329.5	59.1	1541	11 R07304	IgA1 protease. Ha
8	128.5	23.0	1468	20 Y38825	Neisseria gonorrh
9	128	22.9	1431	21 Y75564	Neisseria meningit
10	114	20.4	1394	17 R92768	Adhesion and penet
11	114	20.4	1449	20 Y38824	Neisseria meningit

12	111	19.9	1454	21 Y56621	Neisseria meningit
13	111	19.9	1457	20 Y38823	Neisseria meningit
14	111	19.9	1457	21 B58592	N. meningitidis am
15	111	19.9	1457	21 B25662	N. meningitidis am
16	111	19.9	1457	21 Y56622	Neisseria meningit
17	75.5	13.5	280	21 B63197	Gene 21 human secr
18	75.5	13.5	954	21 B48550	Ruminococcus flave
19	75	13.4	1577	17 R91047	Alpha-D-glucosyltr
20	74.5	13.4	611	17 R88653	N.meningitidis M97
21	74.5	13.4	705	17 R88645	Neisseria meningit
22	73.5	13.2	631	16 R77896	Bacterial transfer
23	73.5	13.2	631	18 W08970	Amino acid sequenc
24	73.5	13.2	631	19 W54128	H. influenzae stra
25	73.5	13.2	631	21 Y51784	H. influenzae non-
26	73.5	13.2	631	21 Y80366	H. influenzae stra
27	71.5	12.8	764	21 B18282	Plasmodium falcipa
28	71	12.7	461	16 R75506	Nicotiana alata ar
29	71	12.7	806	15 R35692	hnRNP U protein.
30	70.5	12.6	1073	21 B01837	Haemophilus influe
31	70.5	12.6	1079	21 B01836	Oyster pearl prote
32	70	12.5	447	17 W01622	Sequence of the As
33	70	12.5	537	7 P60452	Arabidopsis thalia
34	69.5	12.5	115	21 G04598	Arabidopsis thalia
35	68.5	12.3	208	21 G39024	Arabidopsis thalia
36	68.5	12.3	209	21 G36752	Arabidopsis thalia
37	68.5	12.3	226	21 G36751	Arabidopsis thalia
38	68.5	12.3	631	16 R77894	Bacterial transfer
39	68	12.2	405	20 Y34581	Porphorymonas ging
40	68	12.2	461	20 Y34438	Porphorymonas ging
41	68	12.2	1278	21 B18277	Plasmodium falcipa
42	67.5	12.1	631	18 W08968	Amino acid sequenc
43	67.5	12.1	631	19 W54126	H. influenzae stra
44	67.5	12.1	631	21 Y51782	H. influenzae non-
45	67.5	12.1	631	21 Y80364	H. influenzae stra

ALIGNMENTS

RESULT 1
W65656
ID W65656 standard; peptide: 105 AA.
XX
AC W65656;
XX
DT 15-OCT-1998 (first entry)
XX
DE 105-mer peptide used in polysaccharide-peptide conjugate.
XX
KW Polysaccharide-peptide conjugate; vaccine; linker moiety; adjuvant;
KW Immune response.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1 /note= "N-terminal acetyl"
FT
XX
PN W09831393-A2.
XX
PD 23-JUL-1998.
XX
PF 21-JAN-1998; 98WO-EP00654.
XX
XX 21-JAN-1997; 97EP-0100884.
XX
(INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
PI Mistretta N, Moreau M;
XX
DR WPI; 1998-413820/35.
XX
PT Polysaccharide-peptide conjugates, useful as, e.g. vaccines -

PT comprise peptide moiety with at least six amino acid residues,
 PT polysaccharide chain with at least four repeat units, and linker
 PT moiety

XX Example 1; Page 14; 28pp; English.

XX The invention relates to: (A) polysaccharide-peptide conjugate (in which
 CC the polysaccharide is immunogenic), comprising: (a) a peptide moiety
 CC which has at least 6 amino acid residues, at least 1 of which is a
 CC cysteine residue; (b) a polysaccharide chain comprising at least 4 repeat
 CC units, and (c) a linker moiety bound to the thiol group of the cysteine.
 CC The linker is also bound to: (i) native amino, hydroxyl or carboxyl
 CC groups of the polysaccharide chain; (ii) amino groups created by
 CC hydrolysis of native N-acyl groups of the polysaccharide chain, or (iii)
 CC functional groups introduced on the polysaccharide chain upon
 CC derivatisation with a spacer moiety bound to native amino, hydroxyl or
 CC carboxyl groups of the polysaccharide chain, and (B) conjugating a
 CC peptide as in (Aa) to a polysaccharide chain comprising at least 4 repeat
 CC units, comprising: (a) coupling the peptide to a linker through the thiol
 CC group of the cysteine residue, and (b) coupling the linker to the
 CC polysaccharide chain through one of groups (i), (ii) or (iii) as
 CC described in (A). The conjugates are especially useful as vaccines to
 CC elicit a protective long term immune response against a pathogenic
 CC microorganism from which the immunogenic polysaccharide chain is derived.
 CC Known polysaccharide-peptide conjugates are less immunogenic than their
 CC polysaccharide-protein counterparts and require adjuvantation. The new
 CC conjugates have good immunogenicity, largely because the peptides are of
 CC sufficient size. The present sequence represents a peptide used in
 CC a polysaccharide-peptide conjugate.

XX Sequence 105 AA;

Query Match 100.0%; Score 558; DB 19; Length 105;
 Best Local Similarity 100.0%; Pred. No. 1.4e-57;
 Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYYKNRYRYALKSGSVNAPMPENGQTEENNNDWILMGSTQEAAKKNAMHKNQRIISFGSG 60

DB 2 LYYKNRYRYALKSGSVNAPMPENGQTEENNNDWILMGSTQEAAKKNAMHKNQRIISFGSG 61

QY 61 FFGENGKGHNGALNPNFGKSAQNRFLLTGGTNLNGKISVTQG 104

DB 62 ffgengkgghngalnlnfngksagrnflttgntngklsvtqg 105

RESULT 2

W61602
 ID W61602 standard; peptide; 105 AA.

XX AC W61602;

XX DT 27-OCT-1998 (first entry)

XX DE Neisseria IgA1 protease fragment 1.

XX Immunoglobulin protease; carrier; paediatric; vaccine;
 KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.

OS Neisseria sp.

XX WO9831791-A1.

XX PD 23-JUL-1998.

PF 20-JAN-1998; 98WO-EP00294.

XX 21-JAN-1997; 97EP-0100883.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PA (INNR) PASTEUR MERIEUX SERUMS & VACCINS SA.

PI Achtmann M, Moreau M;

XX WPI; 1998-414092/35.

XX New peptide from Neisseria immunoglobulin protease - useful as
 PT immunogenic carrier, e.g. particularly for polysaccharide(s),
 PT forming conjugates used in vaccines against Neisseria and
 PT Haemophilus

PS Claim 6; Page 10; 32pp; English.

XX The Neisseria immunoglobulin protease fragments W61602-W61606 are used
 CC as carriers for a conjugate, particularly in combination with a
 CC polysaccharide. They can be used in paediatric or other vaccines,
 CC particularly for prevention of epidemic bacterial infections, especially
 CC those caused by Neisseria or Haemophilus. The protease fragment is a
 CC highly immunogenic carrier that elicits a T-cell response, resulting in a
 CC long-lasting memory and high antibody titre, and possibly making possible
 CC vaccination without adjuvant.

XX Sequence 105 AA;

Query Match 100.0%; Score 558; DB 19; Length 105;
 Best Local Similarity 100.0%; Pred. NO. 1.4e-57;
 Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYYKNRYRYALKSGSVNAPMPENGQTEENNNDWILMGSTQEAAKKNAMHKNQRIISFGSG 60

DB 2 LYYKNRYRYALKSGSVNAPMPENGQTEENNNDWILMGSTQEAAKKNAMHKNQRIISFGSG 61

QY 61 FFGENGKGHNGALNPNFGKSAQNRFLLTGGTNLNGKISVTQG 104

DB 62 ffgengkgghngalnlnfngksagrnflttgntngklsvtqg 105

RESULT 3

W61603
 ID W61603 standard; peptide; 104 AA.

XX AC W61603;

XX DT 27-OCT-1998 (first entry)

XX DE Neisseria IgA1 protease fragment 2.

XX Immunoglobulin protease; carrier; paediatric; vaccine;
 KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.

OS Neisseria sp.

XX WO9831791-A1.

XX PD 23-JUL-1998.

XX PF 20-JAN-1998; 98WO-EP00294.

XX 21-JAN-1997; 97EP-0100883.

XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PA (INNR) PASTEUR MERIEUX SERUMS & VACCINS SA.

PI Achtmann M, Moreau M;

XX WPI; 1998-414092/35.

XX New peptide from Neisseria immunoglobulin protease - useful as
 PT immunogenic carrier, e.g. particularly for polysaccharide(s),
 PT forming conjugates used in vaccines against Neisseria and
 PT Haemophilus

PS Claim 2; Fig 2; 32pp; English.

XX The Neisseria immunoglobulin protease fragments W61602-W61606 are used

CC as carriers for a conjugate, particularly in combination with a
 CC polysaccharide. They can be used in paediatric or other vaccines,
 CC particularly for prevention of epidemic bacterial infections, especially
 CC those caused by *Neisseria* or *Haemophilus*. The protease fragment is a
 CC highly immunogenic carrier that elicits a T-cell response, resulting in a
 CC long-lasting memory and high antibody titre, and possibly making possible
 CC vaccination without adjuvant.

XX Sequence 104 AA;

Query Match 96.8%; Score 540; DB 19; Length 104;
 Best Local Similarity 96.2%; Pred. No. 1.7e-55;
 Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKKNRYVALKSGGSVNPMPNGQTNNDWILMGSTQEEAKKNAMHNKNNORISGFSG 60
 |||||
 Db 1 lyyknyryyalksggsvnampngvtcnndwfmgytqeaknamhknndgrigsfsg 60

QY 61 FFGKNGKNGHNGALNLFNGKSAQNRFLTGGTNLNGKISVTQ 104
 |||||
 Db 61 ffgkngkghngalnlnfngksaqrflttgtnlngkisvtqg 104

RESULT 4

W61604
 ID W61604 standard; peptide; 104 AA.

AC W61604;

XX 27-OCT-1998 (first entry)

XX *Neisseria* IgA1 protease fragment 3.

XX Immunoglobulin protease; carrier; paediatric; vaccine;
 KW epidemic bacterial infection; *Neisseria*; *Haemophilus*; T-cell.

XX *Neisseria* sp.

OS WO9831791-AL.

PN 23-JUL-1998.

XX 20-JAN-1998; 98WO-EP00294.

XX 21-JAN-1997; 97EP-0100883.

XX (PLAC) MAX PLANCK GES. FOERDERUNG WISSENSCHAFTEN.
 PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX Achtmann M, Moreau M;

XX WPI; 1998-414092/35.

XX New peptide from *Neisseria* immunoglobulin protease - useful as
 PT immunogenic carrier, e.g. particularly for polysaccharide(s),
 PT forming conjugates used in vaccines against *Neisseria* and
 PT *Haemophilus*

PS Claim 2; Fig 1; 32pp; English.

XX The *Neisseria* immunoglobulin protease fragments W61602-W61606 are used
 CC as carriers for a conjugate, particularly in combination with a
 CC polysaccharide. They can be used in paediatric or other vaccines,
 CC particularly for prevention of epidemic bacterial infections, especially
 CC those caused by *Neisseria* or *Haemophilus*. The protease fragment is a
 CC highly immunogenic carrier that elicits a T-cell response, resulting in a
 CC long-lasting memory and high antibody titre, and possibly making possible
 CC vaccination without adjuvant.

XX Sequence 104 AA;

Query Match 96.2%; Score 537; DB 19; Length 104;
 Best Local Similarity 95.2%; Pred. No. 3.8e-55;
 Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKKNRYVALKSGGSVNPMPNGQTNNDWILMGSTQEEAKKNAMHNKNNORISGFSG 60
 |||||
 Db 1 lyyknyryyalksggsvnampngqtenndwfmgytqeaknamhknndgrigsfsg 60

QY 61 FFGKNGKNGHNGALNLFNGKSAQNRFLTGGTNLNGKISVTQ 104
 |||||

Db 61 ffgkngkghngalnlnfngksaqrflttgtnlngkisvtqg 104

RESULT 5

W61606

ID W61606 standard; peptide; 104 AA.

XX AC W61606;

XX 27-OCT-1998 (first entry)

XX *Neisseria* IgA1 protease fragment 5.

XX Immunoglobulin protease; carrier; paediatric; vaccine;
 KW epidemic bacterial infection; *Neisseria*; *Haemophilus*; T-cell.

XX *Neisseria* sp.

XX WO9831791-AL.

XX 23-JUL-1998.

XX 20-JAN-1998; 98WO-EP00294.

XX 21-JAN-1997; 97EP-0100883.

XX (PLAC) MAX PLANCK GES. FOERDERUNG WISSENSCHAFTEN.
 PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX Achtmann M, Moreau M;

XX WPI; 1998-414092/35.

XX New peptide from *Neisseria* immunoglobulin protease - useful as
 PT immunogenic carrier, e.g. particularly for polysaccharide(s),
 PT forming conjugates used in vaccines against *Neisseria* and
 PT *Haemophilus*

PS Claim 2; Fig 4; 32pp; English.

XX The *Neisseria* immunoglobulin protease fragments W61602-W61606 are used
 CC as carriers for a conjugate, particularly in combination with a
 CC polysaccharide. They can be used in paediatric or other vaccines,
 CC particularly for prevention of epidemic bacterial infections, especially
 CC those caused by *Neisseria* or *Haemophilus*. The protease fragment is a
 CC highly immunogenic carrier that elicits a T-cell response, resulting in a
 CC long-lasting memory and high antibody titre, and possibly making possible
 CC vaccination without adjuvant.

XX Sequence 104 AA;

Query Match 88.5%; Score 494; DB 19; Length 104;
 Best Local Similarity 87.5%; Pred. No. 3.9e-50;
 Matches 91; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 LYKKNRYVALKSGGSVNPMPNGQTNNDWILMGSTQEEAKKNAMHNKNNORISGFSG 60
 |||||
 Db 1 lyyknyryyalksggsgrlnampngvaenndwfmgytqeaknamhknndgrigsfsg 60

QY 61 FFGKNGKNGHNGALNLFNGKSAQNRFLTGGTNLNGKISVTQ 104
 |||||

Db 61 ffgkngkghngalnlnfngksaqrflttgtnlngkisvtqg 104

seq. 1

RESULT 6
W61605.
ID W61605 standard; peptide; 104 AA.
XX AC
XX W61605;
XX
DT 27-OCT-1998 (first entry)
XX
DE Neisseria IgA1 protease fragment 4.
XX
XX Immunoglobulin protease; carrier; paediatric; vaccine;
KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.
XX
XX Neisseria sp.
OS
XX WO9831791-A1.
PN
XX 23-JUL-1998.
PD
XX 20-JAN-1998; 98WO-EP00294.
PF
XX 21-JAN-1997; 97EP-0100883.
XX
XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
XX
XX Achtmann M, Moreau M;
PI
XX WPI; 1998-414092/35.
DR
XX
XX New peptide from Neisseria immunoglobulin protease - useful as
PT immunogenic carrier, e.g. particularly for polysaccharide(s),
PT forming conjugates used in vaccines against Neisseria and
PT Haemophilus
XX
XX Claim 2; Fig 1/4; 32pp; English.
PS
XX The Neisseria immunoglobulin protease fragments W61602-W61606 are used
CC as carriers for a conjugate, particularly in combination with a
CC polysaccharide. They can be used in paediatric or other vaccines,
CC particularly for prevention of epidemic bacterial infections, especially
CC those caused by Neisseria or Haemophilus. The protease fragment is a
CC highly immunogenic carrier that elicits a T-cell response, resulting in a
CC long-lasting memory and high antibody titre, and possibly making possible
CC vaccination without adjuvant.
XX
XX Sequence 104 AA;
SQ
Query Match 81.28; Score 453; DB 19; Length 104;
Best Local Similarity 84.58; Pred. No. 2.3e-45;
Matches 82; Conservative 5; Mismatches 10; Indels 0; Gaps 0;
QY 1 LYKKNRYVYALKSGSVNAPMPENGQTNNDWILMGSTQEEAKKNAMNKNQRIISGFSG 60
DB 1 lykknryvayksgsrlnapmpengvaendwvfmgvtqearknammknrrigdfg 60
QY 61 PFGEENGKGHGALNLPNGKSAQNRFLLTGTGNTLNG 97
DB 61 ffdengkgghglnlnfnngksaqnrflttggaunlg 97
RESULT 7
R07304
ID R07304 standard; protein; 1541 AA.
XX
XX R07304;
XX
DT 31-JAN-1991 (first entry)
XX
DE IgA1 protease.

XX IgA1; vaccine; meningitis; gonorrhoea; allergies.
KW
XX Haemophilus influenzae.
OS
XX WO9011367-A.
PN
XX 04-OCT-1990.
PD
XX 16-MAR-1990; 90WO-DK00073.
PF
XX 17-MAR-1989; 89DK-0001308.
XX
XX (KILI/) KILIAN M.
PA
XX Kilian M, Poulsen K;
PI
XX WPI; 1990-320267/42.
DR
XX N-PSDB; O06164.
DR
XX Immunoglobulin A1 protease prodn. - by cloning from
PT microorganisms for immunisation against immunoglobulin A1
PT protease producing bacteria
XX
XX Disclosure; fig 3; 44pp; English.
PS
XX This immunoglobulin (IgA1) protease is produced by recombinant DNA
CC methods. It is useful in a vaccine for e.g. meningococcal meningitis,
CC gonorrhoea or allergic diseases. It specifically cleaves the heavy
CC chain of human IgA1 in the hinge region.
CC
XX Sequence 1541 AA;
SQ
Query Match 59.18; Score 329.5; DB 11; Length 1541;
Best Local Similarity 56.78; Pred. No. 2e-29;
Matches 59; Conservative 19; Mismatches 25; Indels 1; Gaps 1;
QY 1 LYKKNRYVYALKSGSVNAPMPENGQTNNDWILMGSTQEEAKKNAMNKNQRIISGFSG 60
DB 595 lnlentyyalkrgastrelpkngsgesnenwlymgktsdeakrnvmhlnnermngfng 654
QY 61 PFGEENGKGHGALNLPNGKSAQNRFLLTGTGNTLNGKISVTQG 104
DB 655 yfgeegk-nnglnvtfkqseqnrflttgtgntlngdltvekg 697
RESULT 8
Y38825
ID Y38825 standard; Protein; 1468 AA.
XX
XX Y38825;
XX
DT 08-OCT-1999 (first entry)
XX
DE Neisseria gonorrhoeae antigenic protein encoded by ORF1.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicaemia; gonorrhea.
KW
XX Neisseria gonorrhoeae.
OS
XX WO9924578-A2.
PN
XX 20-MAY-1999.
PD
XX 09-OCT-1998; 98WO-IB01665.
PF
XX 01-SEP-1998; 98GB-0019016.
PR 06-NOV-1997; 97GB-0023516.
PR 14-NOV-1997; 97GB-0024190.
PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.

```

PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
PA (CHIR-) CHIRON SPA.
XX
PI Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX WPI; 1999-327407/27.
DR N-PSDB; Z12253.
XX
PT Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection
XX
PS Claim 4; Page 371; 524pp; English.
XX
XX Amino acid sequences Y38499-Y38944 represent Neisseria meningitidis
CC and N. gonorrhoeae antigenic proteins. They are encoded by open
CC reading frames (ORFs) Z11972-Z12358. The antigenic proteins,
CC their fragments, their nucleic acids and antibodies are used for
CC diagnosis, prevention (as vaccines) or treatment of Neisseria
CC infections, such as meningitis, septicemia and gonorrhea. Both
CC organisms are closely related. Fragments of the nucleic acids
CC are useful as hybridisation probes and antisense reagents.
XX
SQ Sequence 1468 AA;

Query Match 23.0%; Score 128.5; DB 20; Length 1468;
Best Local Similarity 35.9%; Pred. No. 4.7e-06;
Matches 28; Conservative 15; Mismatches 26; Indels 9; Gaps 1;

QY 27 TENNDWILMGSTQBEAKKNNHKNRISGFGFFGKNGHNGALNLFNFKSAQNR 86
| | | | | : : : : : : : : : : : : : : : : : : : : : :
| | | | | : : : : : : : : : : : : : : : : : : : : : :
Db 570 tgnkdittgnn-----nnldskkeiayngwfgekdatktnglnlnyppeadrt 620

QY 87 FLITGGTNLNGKISVTQG 104
| | | | | : : : : : : : : : : : : : : : : : : : : : :
| | | | | : : : : : : : : : : : : : : : : : : : : : :
Db 621 lllsggtnglnqntqng 638

RESULT 9
Y75564
ID Y75564 standard; Protein; 1431 AA.
XX
AC Y75564;
XX
DT 21-MAR-2000 (first entry)
XX
XX Neisseria meningitidis ORF 759 protein sequence SEQ ID NO:2602.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicemia;
KW antibacterial; gene therapy.
XX
XX Neisseria meningitidis.
XX
OS
XX
XX WO9957280-A2.
XX
PD 11-NOV-1999.
XX
XX 30-APR-1999; 99WO-US09346.
XX
XX 01-MAY-1998; 98US-0083758.
XX 31-JUL-1998; 98US-0094869.
XX 02-SEP-1998; 98US-0098994.
XX 02-SEP-1998; 98US-0099062.
XX 09-OCT-1998; 98US-0103749.
XX 09-OCT-1998; 98US-0103794.
XX 09-OCT-1998; 98US-0103796.
XX 25-FEB-1999; 99US-0121528.
XX
XX (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.

```

```

XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
XX WPI; 2000-062150/05.
DR N-PSDB; Z54326.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics
XX
XX Claim 2; Page 1234; 1453pp; English.
XX
XX Z53015 to Z54536, Z54577 to Z54615, and Y74253 to Y75941 represent
CC novel Neisseria meningitis and N. gonorrhoeae polynucleotides and
CC polypeptides. Z54537 to Z54576 and Z54616 to Z55473 represent PCR
CC primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 1431 AA;

Query Match 22.9%; Score 128; DB 21; Length 1431;
Best Local Similarity 25.5%; Pred. No. 5.2e-06;
Matches 35; Conservative 22; Mismatches 44; Indels 36; Gaps 5;

QY 3 YKNVR--YYALKGGGVNAMPENGTENNNDWILMGSTQBEAKKNNHKN--NORISGFS 59
| | | | | : : : : : : : : : : : : : : : : : : : : : :
| | | | | : : : : : : : : : : : : : : : : : : : : : :
Db 555 hnrtrtdyflkpggnpreffpln-mknstswqfignrggaeqvaaenarpdlitfg 613

QY 60 GFEGE--ENGKG-----HNGALNLFNFKSAQNR 87
| | | | | : : : : : : : : : : : : : : : : : : : : : :
| | | | | : : : : : : : : : : : : : : : : : : : : : :
Db 614 gylgenadtgaapsysktnaealektrhianaavryrpyryngalnhyrpkrtdstl 673

QY 88 LLTGGTNLNGKISVTQG 104
| | | | | : : : : : : : : : : : : : : : : : : : : : :
| | | | | : : : : : : : : : : : : : : : : : : : : : :
Db 674 lllggmnlngvliegg 690

RESULT 10
R92768
ID R92768 standard; Protein; 1394 AA.
XX
AC R92768;
XX
XX 01-JUN-1996 (first entry)
XX
XX Adhesion and penetration protein.
XX
XX Adhesion and penetration protein; hap gene; signal peptide;
KW protease; helper domain; pore; outer membrane protein; beta-domain;
KW secretion; active site; catalytic domain; recombinant vaccine;
KW monoclonal antibody; diagnostic; immunoassay.
XX
XX Haemophilus influenzae.
XX
XX Key Location/Qualifiers
XX Peptide 1..25
XX /note= "Signal peptide"
XX Peptide 27..33
XX /note= "Mature protein N-terminal peptide"
XX Protein 26..959
XX /note= "Secreted 110-kDa protease fragment"
XX Domain 241..248
XX /note= "Putative catalytic domain"

```

```

FT Active-site 243 /note= "Active site serine residue"
FT Region 698 /note= "Conserved cysteine residue"
FT Region 709 /note= "Conserved cysteine residue"
FT Cleavage-site 960 /note= "Site for cleavage of secreted protease from
FT outer membrane protein"
FT Protein 960..1394
FT /note= "45-kDa outer membrane protein fragment"
XX
XX WO9603858-A1.
XX
XX 29-FEB-1996.
XX
XX 16-AUG-1995; 95WO-US10661.
XX
XX 25-AUG-1994; 94US-0296791.
XX
XX (STRD ) UNIV LELAND STANFORD JUNIOR.
XX (UNIW ) UNIV WASHINGTON.
XX
XX Falkow S, St Geme JW;
XX
XX WPI; 1996-151147/15.
XX
XX N-PSDB; T17215.
XX
XX Haemophilus adhesion and penetration protein and corresponding DNA
XX - used to produce vaccines against H. influenzae infection
XX
XX Claim 2; Fig 6; 105pp; English.
XX
XX The sequence represents a Haemophilus influenzae adhesion and
XX penetration protein, encoded by the hap gene. The protein is
XX first synthesised as a preprotein with 3 functional domains (the
XX N-terminal signal peptide, the protease and a C-terminal helper
XX domain), which is transported to the periplasm, followed by
XX insertion of the C-terminal beta-domain into the outer membrane,
XX possibly forming a pore, and export of the N-terminal fragment
XX through the outer membrane, followed by autolytic cleavage
XX and secretion of the mature protease, leaving an outer membrane
XX protein fragment. An active site serine residue characteristic of
XX serine proteases is present, along with 2 cysteine residues which
XX may be important in forming the catalytic domain. A gene encoding
XX the protein may be inserted in a vector and expressed in
XX recombinant host cells, for use as a recombinant vaccine. The
XX protein may also be used in diagnostic monoclonal antibody production.
XX
XX Sequence 1394 AA;

Query Match 20.4%; Score 114; DB 17; Length 1394;
Best Local Similarity 38.3%; Pred. No. 0.00021;
Matches 23; Conservative 13; Mismatches 24; Indels 0; Gaps 0;

QY 45 NAMHNKNNQISGSGFFGEGNGKNGALNLFNGKSAQNRFLLTGTNLNGKISVTQ 104
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 551 nninkldyrkeiayngwfgedkknkngnllykpttedrtlllsggtlkgdltqtkg 610

RESULT 11
Y38824
ID Y38824 standard; Protein; 1449 AA.
XX
XX Y38824;
AC
XX
XX 08-OCT-1999 (first entry)
DT
XX
XX Neisseria meningitidis strain A antigen encoded by ORF1.
DE
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW
XX treatment; Neisseria infection; meningitis; septicaemia; gonorrhea.
KW

```

```

XX
OS Neisseria meningitidis.
XX
XX WO9924578-A2.
XX
XX 20-MAY-1999.
XX
XX 09-OCT-1998; 98WO-IB01665.
XX
XX 01-SEP-1998; 98GB-0019016.
XX 06-NOV-1997; 97GB-0023516.
XX 14-NOV-1997; 97GB-0024190.
XX 18-NOV-1997; 97GB-0024386.
XX 27-NOV-1997; 97GB-0025158.
XX 10-DEC-1997; 97GB-0026147.
XX 14-JAN-1998; 98GB-0000759.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Grandi G, Masignani V, Pizzo M, Rappuoli R, Scarlato V;
XX
XX WPI; 1999-327407/27.
XX N-PSDB; Z12252.
XX
XX Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
XX diagnosis, treatment and prevention of infection
XX
XX Claim 4; Page 365; 524pp; English.
XX
XX Amino acid sequences Y38499-Y38944 represent Neisseria meningitidis
XX and N. gonorrhoeae antigenic proteins. They are encoded by open
XX reading frames (ORFs) Z11972-Z12358. The antigenic proteins,
XX their fragments, their nucleic acids and antibodies are used for
XX diagnosis, prevention (as vaccines) or treatment of Neisseria
XX infections, such as meningitis, septicaemia and gonorrhea. Both
XX organisms are closely related. Fragments of the nucleic acids
XX are useful as hybridisation probes and antisense reagents.
XX
XX Sequence 1449 AA;

Query Match 20.4%; Score 114; DB 20; Length 1449;
Best Local Similarity 38.8%; Pred. No. 0.00022;
Matches 26; Conservative 11; Mismatches 28; Indels 2; Gaps 1;

QY 38 TOEAKKNNHKNQISGSGFFGEGNGKNGALNLFNGKSAQNRFLLTGTNLNG 97
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 567 tpsgkn--inrlnyskeiayngwfgedkttktnglrlvygpaadrtxlsggtning 624

QY 98 KISVTQ 104
| : | |
Db 625 nitqtng 631

RESULT 12
Y56621
ID Y56621 standard; Protein; 1454 AA.
XX
XX AC Y56621;
XX
XX 21-FEB-2000 (first entry)
DT
XX
XX Neisseria meningitidis strain ATCC13090 BASB006 protein sequence.
DE
XX
XX Neisseria meningitidis; BASB006; diagnosis; bacterial; infection;
XX vaccine; antibiotic; upper respiratory tract infection; meningitis;
XX invasive bacterial disease; bacteraemia; screening; antibacterial.
XX
XX Neisseria meningitidis.
OS
XX
XX WO9955873-A2.
XX
XX 04-NOV-1999.
XX

```



```

XX (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
XX Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;
XX
XX WPI: 2000-647603/62.
DR N-PSDB; F21587.
XX
XX Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections -
XX
XX Example 1; Page 114; 692pp; English.
XX
XX The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in F21544 and F21607 to
CC F21613 represent fragments of the NMB genomic sequence, as the sequence
CC was too long to go in a record on its own it was split into 8 sequences
CC which overlap each other at the beginning and end of each sequence by
CC 4980 bp (i.e. the last 4980 bp of F21544 is repeated at the beginning
CC of F21607, the last 4980 bp of F21607 are repeated at the beginning of
CC F21608, and so on). F21545 to F21588 encode the Neisseria proteins given
CC in B58550 to B58593, and F21589 to F21606 represent PCR primers which
CC are used in the exemplification of the present invention. The NMB genome
CC and fragments from it have antibacterial activity, and can be used in
CC vaccines and gene therapy. Neisseria nucleic acids, proteins and/or
CC antibodies which binds to the proteins can be used in compositions for
CC treating or preventing infection due to Neisserial bacteria or as a
CC diagnostic reagent for detecting the presence of Neisserial bacteria or
CC of antibodies raised to Neisserial bacteria. Computers, computer memory,
CC computer storage medium or computer databases can be used in a search to
CC identify open reading frames (ORFs) or coding sequences within the NMB
CC genome. The DNA sequences provide further opportunities to find
CC antigenic or immunogenic proteins which are more effective in vaccines
XX
XX Sequence 1457 AA;
XX
XX Query Match 19.9%; Score 111; DB 21; Length 1457;
XX Best Local Similarity 29.6%; Pred. No. 0.00051;
XX Matches 24; Conservative 18; Mismatches 39; Indels 0; Gaps 0;
XX
Oy 24 NCQTEENDWILMGSTQBEAKKNAMHNKNRISGFSGFFGKNGKGNALNFKGSA 83
Db 558 nhngdkestvltgkdiattgnnsidskkelayngwfgkdkttkngnlvlyqpaae 617
XX
Oy 84 QNRFLTGCTNLNGKISVTQG 104
Db 618 drtllsggtlnngnitqng 638
XX
RESULT 15
B25662
ID B25662 standard; Protein; 1457 AA.
XX
XX B25662;
XX
XX 04-DEC-2000 (first entry)
XX
XX N. meningitidis amino acid sequence orfl-1.ppe SEQ ID NO:1047.
DE
XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB.
XX
XX Neisseria meningitidis.
OS
XX WO200022430-A2.
XX
XX 20-APR-2000.
XX

```

```

PF 08-OCT-1999; 99WO-US23573.
XX
XX 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
XX (CHIR ) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
XX WPI: 2000-318079/27.
DR N-PSDB; A81302.
XX
XX Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
XX other Neisserial infections, for example, N.gonorrhoea -
XX
XX Example 1; Page 113; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic
XX proteins from Neisseria genomic sequences. A81453 to A82414 represent
XX specifically claimed Neisseria meningitidis genomic DNA sequences; A81260
XX to A81303 and B25620 to B25663 represent Neisseria DNA sequences and
XX their corresponding proteins; A81254 to A81259 and A81304 to A81321
XX represent PCR primers used in the isolation of Neisseria meningitidis DNA
XX sequences; and A81322 to A81452 represent Neisseria meningitidis MenB
XX polynucleotide ORF sequences, which are all used in the exemplification
XX of the present invention. The nucleic acid sequences, protein sequences,
XX and antibodies against them, can be used in the manufacture of a
XX composition. The composition can be used as a medicament (or in the
XX manufacture of a medicament) for treating, preventing or diagnosing
XX infection due to Neisserial bacteria. For example, some of the identified
XX proteins could be components of vaccines against Meningococcus B; against
XX all serotypes; and/or against all pathogenic Neisseriae. Identification
XX of sequences from the bacterium will also facilitate production of
XX biological probes, particularly organism-specific probes. Attempts to
XX make efficacious Meningococcus B vaccines have failed mainly due to
XX antigen tolerance. Multivalent vaccines have also been tried but none
XX have successfully overcome antigenic variability. The provision of
XX further, complete sequences may provide an opportunity to identify
XX secreted or surface exposed proteins that may be presumed targets for the
XX immune system and which are not antigenically variable or at least more
XX conserved than other more variable regions.
XX
XX Sequence 1457 AA;
XX
XX Query Match 19.9%; Score 111; DB 21; Length 1457;
XX Best Local Similarity 29.6%; Pred. No. 0.00051;
XX Matches 24; Conservative 18; Mismatches 39; Indels 0; Gaps 0;
XX
Oy 24 NCQTEENDWILMGSTQBEAKKNAMHNKNRISGFSGFFGKNGKGNALNFKGSA 83
Db 558 nhngdkestvltgkdiattgnnsidskkelayngwfgkdkttkngnlvlyqpaae 617
XX
Oy 84 QNRFLTGCTNLNGKISVTQG 104
Db 618 drtllsggtlnngnitqng 638
XX
Search completed: May 1, 2001, 14:58:47
Job time: 609 sec

```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 14:58:47 ; Search time 66.74 Seconds
(without alignments)
89.077 Million cell updates/sec

Title: US-09-142-970-2
Perfect score: 562
Sequence: 1 LYNNRYALKSGSVNAP.....NRLLTGTGTLNGKISVTQG 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	562	100.0	104	19 W61603	Neisseria IgA1 pro
2	546	97.2	104	19 W61604	Neisseria IgA1 pro
3	540	96.1	105	19 W65656	105-mer peptide us
4	540	96.1	105	19 W61602	Neisseria IgA1 pro
5	514	91.5	104	19 W61606	Neisseria IgA1 pro
6	475	84.5	104	19 W61605	Neisseria IgA1 pro
7	329.5	58.6	154.1	11 R07304	IgA1 protease. Ha
8	133	23.7	143.1	21 Y75564	Neisseria meningit
9	128.5	22.9	146.8	20 Y38825	Neisseria gonorrhoe
10	114	20.3	139.4	17 R92768	Adhesion and penet
11	114	20.3	144.9	20 Y38824	Neisseria meningit

12	111.5	19.8	1454	21 Y56621	Neisseria meningit
13	111.5	19.8	1457	20 Y38823	Neisseria meningit
14	111.5	19.8	1457	21 B58592	N. meningitidis am
15	111.5	19.8	1457	21 B25662	N. meningitidis am
16	111.5	19.8	1457	21 Y56622	Neisseria meningit
17	83	14.8	1577	17 R91047	Alpha-D-glucosyltr
18	79.5	14.1	611	17 R88653	N. meningitidis M97
19	79.5	14.1	705	17 R88645	Neisseria meningit
20	78.5	14.0	1073	21 B01837	Haemophilus influe
21	78.5	14.0	1079	21 B01836	Haemophilus influe
22	77.5	13.8	631	16 R77894	Bacterial transfer
23	76.5	13.6	631	18 W08968	Amino acid sequenc
24	76.5	13.6	631	19 W54126	H. influenzae stra
25	76.5	13.6	631	21 Y51782	H. influenzae non-
26	76.5	13.6	631	21 Y80364	H. influenzae stra
27	76	13.5	461	16 R75506	Nicotiana alata ar
28	75.5	13.4	764	21 B18282	Plasmodium falcipa
29	74.5	13.3	631	16 R77896	Bacterial transfer
30	74.5	13.3	631	18 W08970	Amino acid sequenc
31	74.5	13.3	631	19 W54128	H. influenzae stra
32	74.5	13.3	631	21 Y51784	H. influenzae non-
33	74.5	13.3	631	21 Y80366	H. influenzae stra
34	73.5	13.1	671	17 R85290	Streptococcus faec
35	72.5	12.9	447	17 W01622	Oyster pearl prote
36	71.5	12.7	115	21 G04598	Arabidopsis thalia
37	71	12.6	806	15 R55692	hnRNP U protein.
38	70	12.5	275	21 G57194	Arabidopsis thalia
39	70	12.5	288	21 G57193	Arabidopsis thalia
40	70	12.5	311	21 G57192	Arabidopsis thalia
41	69.5	12.4	1247	21 B18215	Plasmodium falcipa
42	68.5	12.2	280	21 B63197	Gene 21 human secr
43	68.5	12.2	338	18 W22503	Ubiquitin-ribosoma
44	68.5	12.2	437	16 R75507	Nicotiana plumbagi
45	68.5	12.2	954	21 B48550	Ruminococcus flave

ALIGNMENTS

RESULT 1
W61603
ID W61603 standard; peptide: 104 AA.
XX AC
XX W61603;
XX
DT 27-OCT-1998 (first entry)
XX
DE Neisseria IgA1 protease fragment 2.
XX
KW Immunoglobulin protease; carrier; paediatric; vaccine;
KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.
XX
OS Neisseria sp.
PN W09831791-A1.
XX
PD 23-JUL-1998.
XX
PF 20-JAN-1998; 98WO-EP00294.
XX
PR 21-JAN-1997; 97EP-0100883.
XX
(PLAC) MAX PLANCK GFS FOERDERUNG WISSENSCHAFTEN.
(INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

Achtmann M, Moreau M;
WPI: 1998-414092/35.

New peptide from Neisseria immunoglobulin protease - useful as immunogenic carrier, e.g. particularly for polysaccharide(s), forming conjugates used in vaccines against Neisseria and Haemophilus

Applicant

```
XX PS Claim 2; Fig 2; 32pp; English.
XX CC The Neisseria immunoglobulin protease fragments W61602-W61606 are used
XX CC as carriers for a conjugate, particularly in combination with a
XX CC polysaccharide. They can be used in paediatric or other vaccines,
XX CC particularly for prevention of epidemic bacterial infections, especially
XX CC those caused by Neisseria or Haemophilus. The protease fragment is a
XX CC highly immunogenic carrier that elicits a T-cell response, resulting in a
XX CC long-lasting memory and high antibody titre, and possibly making possible
XX CC vaccination without adjuvant.
XX SQ Sequence 104 AA;

Query Match 100.0%; Score 562; DB 19; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.5e-57;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYYKNRYALKSGSVNAPMPENGVTENNDFMGYTQEEAKKNAMNHNKNNQISGFSG 60
DB 1 lyyknyryalksgsvnapmpengvtenndwvfmgytqeeakknammhknqrisgfsq 60
QY 61 FFGENGKGHNGALNLFNGKSAQNRFLLTGTNLNGKISVTQG 104
DB 61 ffeengkgghngalnlnfngksaqrfltlgtnglnkgsivtqg 104

RESULT 2
W61604
ID W61604 standard; peptide; 104 AA.
XX AC W61604;
XX DT 27-OCT-1998 (first entry)
XX DE Neisseria IgA1 protease fragment 3.
XX KW Immunoglobulin protease; carrier; paediatric; vaccine;
XX KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.
XX OS Neisseria sp.
XX PN W09831791-Al.
XX PD 23-JUL-1998.
XX PF 20-JAN-1998; 98WO-EP00294.
XX PR 21-JAN-1997; 97EP-0100883.
XX PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX PA (INNR ) PASTEUR MERIEUX SERUMS & VACCINS SA.
XX PI Achtmann M, Moreau M;
XX PS WPI; 1998-414092/35.
XX DR
XX PT New peptide from Neisseria immunoglobulin protease - useful as
XX PT immunogenic carrier, e.g. particularly for polysaccharide(s),
XX PT forming conjugates used in vaccines against Neisseria and
XX PT Haemophilus
XX PS Claim 2; Fig 1; 32pp; English.
XX CC The Neisseria immunoglobulin protease fragments W61602-W61606 are used
XX CC as carriers for a conjugate, particularly in combination with a
XX CC polysaccharide. They can be used in paediatric or other vaccines,
XX CC particularly for prevention of epidemic bacterial infections, especially
XX CC those caused by Neisseria or Haemophilus. The protease fragment is a
XX CC highly immunogenic carrier that elicits a T-cell response, resulting in a
XX CC long-lasting memory and high antibody titre, and possibly making possible
XX CC vaccination without adjuvant.

XX SQ Sequence 104 AA;

Query Match 97.2%; Score 546; DB 19; Length 104;
Best Local Similarity 97.1%; Pred. No. 1.1e-55;
Matches 101; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LYYKNRYALKSGSVNAPMPENGVTENNDFMGYTQEEAKKNAMNHNKNNQISGFSG 60
DB 1 lyyknyryalksgsvnapmpengvtenndwvfmgytqeeakknammhknqrisgfsq 60
QY 61 FFGENGKGHNGALNLFNGKSAQNRFLLTGTNLNGKISVTQG 104
DB 61 ffeengkgghngalnlnfngksaqrfltlgtnglnkgsivtqg 104

RESULT 3
W65656
ID W65656 standard; peptide; 105 AA.
XX AC W65656;
XX DT 15-OCT-1998 (first entry)
XX DE 105-mer peptide used in polysaccharide-peptide conjugate.
XX KW Polysaccharide-peptide conjugate; vaccine; linker moiety; adjuvant;
XX KW immune response.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note="N-terminal acetyl"
XX ET
XX PN W09831393-A2.
XX PD 23-JUL-1998.
XX PF 21-JAN-1998; 98WO-EP00654.
XX PR 21-JAN-1997; 97EP-0100884.
XX PA (INNR ) PASTEUR MERIEUX SERUMS & VACCINS SA.
XX PI Mistretta N, Moreau M;
XX PS WPI; 1998-413820/35.
XX DR
XX PT Polysaccharide-peptide conjugates, useful as, e.g. vaccines -
XX PT comprise peptide moiety with at least six amino acid residues,
XX PT polysaccharide chain with at least four repeat units, and linker
XX PT moiety
XX PS Example 1; Page 14; 28pp; English.
XX CC The invention relates to: (A) polysaccharide-peptide conjugate (in which
XX CC the polysaccharide is immunogenic), comprising: (a) a peptide moiety
XX CC which has at least 6 amino acid residues, at least 1 of which is a
XX CC cysteine residue; (b) a polysaccharide chain comprising at least 4 repeat
XX CC units, and (c) a linker moiety bound to the thiol group of the cysteine.
XX CC The linker is also bound to: (i) native amino, hydroxyl or carboxyl
XX CC groups of the polysaccharide chain; (ii) amino groups created by
XX CC hydrolysis of native N-acyl groups of the polysaccharide chain, or (iii)
XX CC functional groups introduced on the polysaccharide chain upon
XX CC derivatisation with a spacer moiety bound to native amino, hydroxyl or
XX CC carboxyl groups of the polysaccharide chain, and (B) conjugating a
XX CC peptide as in (Aa) to a polysaccharide chain comprising at least 4 repeat
XX CC units, comprising: (a) coupling the peptide to a linker through the thiol
XX CC group of the cysteine residue, and (b) coupling the linker to the
XX CC polysaccharide chain through one of groups (i), (ii) or (iii) as
XX CC described in (A). The conjugates are especially useful as vaccines to
```

CC elicit a protective long term immune response against a pathogenic
 CC microorganism from which the immunogenic polysaccharide chain is derived.
 CC known polysaccharide-peptide conjugates are less immunogenic than their
 CC polysaccharide-protein counterparts and require adjuvantation. The new
 CC conjugates have good immunogenicity, largely because the peptides are of
 CC sufficient size. The present sequence represents a peptide used in
 CC a polysaccharide-peptide conjugate.
 XX
 SQ Sequence 105 AA;

Query Match 96.1%; Score 540; DB 19; Length 105;
 Best Local Similarity 96.2%; Pred. No. 5.3e-55;
 Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKYNRYALKSGSVNAPMPENGVTENNNDWVFMGYTQEEAKKNAMHNKNNORISGFSG 60
 |||||
 Db 2 LYKYNRYALKSGSVNAPMPENGVTENNNDWVFMGYTQEEAKKNAMHNKNNORISGFSG 61
 |||||
 QY 61 FFGENGKGHNGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 104
 |||||
 Db 62 ffgengkgngalnlnfngksaqnrflttggtngnkisvtgg 105
 |||||

RESULT 4
 W61602 ID W61602 standard; peptide; 105 AA.
 XX AC W61602;
 XX DT 27-OCT-1998 (first entry)
 XX DE Neisseria IgA1 protease fragment 1.
 XX KW Immunoglobulin protease; carrier; paediatric; vaccine;
 KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.
 XX OS Neisseria sp.
 XX PN WO9831791-AL.
 XX PD 23-JUL-1998.
 XX PF 20-JAN-1998; 98WO-EP00294.
 XX PR 21-JAN-1997; 97EP-0100883.
 XX PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
 XX PI Achtmann M, Moreau M;
 XX WPI; 1998-414092/35.
 XX DR 20-JAN-1998; 98WO-EP00294.
 XX PF 21-JAN-1997; 97EP-0100883.
 XX PR (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
 XX PI Achtmann M, Moreau M;
 XX WPI; 1998-414092/35.
 XX DR New peptide from Neisseria immunoglobulin protease - useful as
 PT immunogenic carrier, e.g. particularly for polysaccharide(s),
 PT forming conjugates used in vaccines against Neisseria and
 PT Haemophilus
 XX
 XX PS Claim 6; Page 10; 32pp; English.
 XX CC The Neisseria immunoglobulin protease fragments W61602-W61606 are used
 CC as carriers for a conjugate, particularly in combination with a
 CC polysaccharide. They can be used in paediatric or other vaccines,
 CC particularly for prevention of epidemic bacterial infections, especially
 CC those caused by Neisseria or Haemophilus. The protease fragment is a
 CC highly immunogenic carrier that elicits a T-cell response, resulting in a
 CC long-lasting memory and high antibody titre, and possibly making possible
 CC vaccination without adjuvant.
 XX
 XX SQ Sequence 105 AA;

Query Match 96.1%; Score 540; DB 19; Length 105;
 Best Local Similarity 96.2%; Pred. No. 5.3e-55;
 Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 1 LYKYNRYALKSGSVNAPMPENGVTENNNDWVFMGYTQEEAKKNAMHNKNNORISGFSG 60
 |||||
 Db 2 LYKYNRYALKSGSVNAPMPENGVTENNNDWVFMGYTQEEAKKNAMHNKNNORISGFSG 61
 |||||
 QY 61 FFGENGKGHNGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 104
 |||||
 Db 62 ffgengkgngalnlnfngksaqnrflttggtngnkisvtgg 105
 |||||

RESULT 5
 W61606 ID W61606 standard; peptide; 104 AA.
 XX AC W61606;
 XX DT 27-OCT-1998 (first entry)
 XX DE Neisseria IgA1 protease fragment 5.
 XX KW Immunoglobulin protease; carrier; paediatric; vaccine;
 KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.
 XX OS Neisseria sp.
 XX PN WO9831791-AL.
 XX PD 23-JUL-1998.
 XX PF 20-JAN-1998; 98WO-EP00294.
 XX PR 21-JAN-1997; 97EP-0100883.
 XX PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
 PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
 XX PI Achtmann M, Moreau M;
 XX WPI; 1998-414092/35.
 XX DR New peptide from Neisseria immunoglobulin protease - useful as
 PT immunogenic carrier, e.g. particularly for polysaccharide(s),
 PT forming conjugates used in vaccines against Neisseria and
 PT Haemophilus
 XX
 XX PS Claim 2; Fig 4; 32pp; English.
 XX CC The Neisseria immunoglobulin protease fragments W61602-W61606 are used
 CC as carriers for a conjugate, particularly in combination with a
 CC polysaccharide. They can be used in paediatric or other vaccines,
 CC particularly for prevention of epidemic bacterial infections, especially
 CC those caused by Neisseria or Haemophilus. The protease fragment is a
 CC highly immunogenic carrier that elicits a T-cell response, resulting in a
 CC long-lasting memory and high antibody titre, and possibly making possible
 CC vaccination without adjuvant.
 XX
 XX SQ Sequence 104 AA;

Query Match 91.5%; Score 514; DB 19; Length 104;
 Best Local Similarity 89.4%; Pred. No. 5.3e-52;
 Matches 93; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 QY 1 LYKYNRYALKSGSVNAPMPENGVTENNNDWVFMGYTQEEAKKNAMHNKNNORISGFSG 60
 |||||
 Db 1 LYKYNRYALKSGSVNAPMPENGVTENNNDWVFMGYTQEEAKKNAMHNKNNORISGFSG 60
 |||||

QY 61 FFGENGKGHNGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 104
 |||||
 Db 61 ffgengkgngalnlnfngksaqnrflttggtngnkisvtgg 104
 |||||

RESULT 6
W61605
ID W61605 standard; peptide; 104 AA.
XX
AC W61605;
XX
DT 27-OCT-1998 (first entry)
XX
DE Neisseria IgA1 protease fragment 4.
XX
KW Immunoglobulin protease; carrier; paediatric; vaccine;
KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.
XX
OS Neisseria sp.
XX
PN W09831791-A1.
XX
PD 23-JUL-1998.
XX
PF 20-JAN-1998; 98WO-EP00294.
XX
PR 21-JAN-1997; 97EP-0100883.
XX
PA (PLAC) MAX PLANCK GES FORDERUNG WISSENSCHAFTEN.
XX (INNR) PASTEUR MERIEUX SERUMS & VACCINS SA.
XX
PI Achtmann M, Moreau M;
XX
XX WPI; 1998-414092/35.
XX
XX New peptide from Neisseria immunoglobulin protease - useful as
PT immunogenic carrier, e.g. particularly for polysaccharide(s),
PT forming conjugates used in vaccines against Neisseria and
PT Haemophilus
XX
PS Claim 2; Fig 1/4; 32pp; English.
XX
XX The Neisseria immunoglobulin protease fragments W61602-W61606 are used
CC as carriers for a conjugate, particularly in combination with a
CC polysaccharide. They can be used in paediatric or other vaccines,
CC particularly for prevention of epidemic bacterial infections, especially
CC those caused by Neisseria or Haemophilus. The protease fragment is a
CC highly immunogenic carrier that elicits a T-cell response, resulting in a
CC long-lasting memory and high antibody titre, and possibly making possible
CC vaccination without adjuvant.
XX
SQ Sequence 104 AA;

Query Match 84.5%; Score 475; DB 19; Length 104;
Best Local Similarity 88.7%; Pred. No. 1.7e-47;
Matches 86; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 LYYKNRYVALKSGSVNAPMPENGVTENNDFVFGYTQEEAKNKNHNNQRISGFSG 60
Db 1 Lyyknryvalksggrlnapmpengvaenndwfmgytqeeaknknhnnrrigdfg 60
QY 61 PFGEENGKNGHNGALNFGKSAQNRFLTGTGTLNG 97
Db 61 fideengkngalnfgksaqnrflitgnganng 97

RESULT 7
R07304
ID R07304 standard; protein; 1541 AA.
XX
AC R07304;
XX
DT 31-JAN-1991 (first entry)
XX
DE IgA1 protease.

XX IgA1; vaccine; meningitis; gonorrhoea; allergies.
XX Haemophilus influenzae.
OS
PN W09011367-A.
XX
PD 04-OCT-1990.
XX
PF 16-MAR-1990; 90WO-DK00073.
XX
PR 17-MAR-1989; 89DK-0001308.
XX (KILI/) KILIAN M.
XX
PI Kilian M, Poulsen K;
XX
DR WPI; 1990-320267/42.
DR N-PSDB; Q06164.
XX
XX Immunoglobulin A1 protease prodn. - by cloning from
PT microorganisms for immunisation against immunoglobulin A1
PT protease producing bacteria
XX
PS Disclosure; fig 3; 44pp; English.
XX
CC This immunoglobulin (Ig)A1 protease is produced by recombinant DNA
CC methods. It is useful in a vaccine for e.g. meningococcal meningitis,
CC gonorrhoea or allergic diseases. It specifically cleaves the heavy
CC chain of human IgA1 in the hinge region.
XX
SQ Sequence 1541 AA;

Query Match 58.6%; Score 329.5; DB 11; Length 1541;
Best Local Similarity 56.7%; Pred. No. 3.9e-29;
Matches 59; Conservative 20; Mismatches 24; Indels 1; Gaps 1;

QY 1 LYYKNRYVALKSGSVNAPMPENGVTENNDFVFGYTQEEAKNKNHNNQRISGFSG 60
Db 595 lnlenyyvalksgatrseplkngsesnenwlymgktsdeakrnvmhnnernmgng 654
QY 61 PFGEENGKNGHNGALNFGKSAQNRFLTGTGTLNGKISVTQG 104
Db 655 yfgeegk-ngnlnvtfkgseqrflitggtngldltvek 697

RESULT 8
Y75564
ID Y75564 standard; Protein; 1431 AA.
XX
AC Y75564;
XX
DT 21-MAR-2000 (first entry)
XX
DE Neisseria meningitidis ORF 759 protein sequence SEQ ID NO:2602.
XX
KW Neisseria meningitidis; Neisseria gonorrhoea; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene/therapy.
XX
OS Neisseria meningitidis.
XX
PN W09957280-A2.
XX
PD 11-NOV-1999.
XX
PF 30-APR-1999; 99WO-US09346.
XX
PR 01-MAY-1998; 98US-0083758.
PR 31-JUL-1998; 98US-0094869.
PR 02-SEP-1998; 98US-0098994.
PR 02-SEP-1998; 98US-0099062.

Seg 10 Nov 2


```

PR 09-OCT-1998; 98US-0103749.
PR 09-OCT-1998; 98US-0103794.
PR 09-OCT-1998; 98US-0103796.
PR 25-FEB-1999; 99US-0121528.
XX
PA (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX
DR WPI; 2000-062150/05.
DR N-PSDB; 254326.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
XX vaccines and diagnostics
XX
PS Claim 2; Page 1234; 1453pp; English.
XX
CC 253015 to 254536, 254577 to 254615, and Y74253 to Y75941 represent
CC novel Neisseria meningitis and N. gonorrhoeae polynucleotides and
CC polypeptides. 254537 to 254576 and 254616 to 255473 represent PCR
CC primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisseria bacteria (e.g. meningitis and septicemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
SQ Sequence 1431 AA;

Query Match 23.7%; Score 133; DB 21; Length 1431;
Best Local Similarity 26.3%; Pred. No. 1.8e-06;
Matches 36; Conservative 22; Mismatches 43; Indels 36; Gaps 5;

QY 3 YKNYR--YYALKSGGSYNAPMPENGVTENNNDWVFMTQEEAKKNAMNHN--NORISGFS 59
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 555 hnrtrdyfllkpggnprfllpn-mknstswqfignrqgaeeqvaqaenrpdllitfg 613
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 60 GFCE--ENGKG-----HNGALNLFNFKSAONRF 87
   | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 614 gylgenagtgaapsysktneaaiekrhianaavgyrpeyrngalnlnhyrpkrtstl 673

QY 88 LLTGGTNLNGKISVTQG 104
   | | | | | | | | : |
Db 674 lllgmgmlngevllieg 690

RESULT 9
Y38825
ID Y38825 standard; Protein; 1468 AA.
XX
AC Y38825;
XX
XX 08-OCT-1999 (first entry)
DT
DE Neisseria gonorrhoeae antigenic protein encoded by ORF1.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW treatment; Neisseria infection; meningitis; septicemia; gonorrhea.
XX
XX Neisseria gonorrhoeae.
OS
XX WO9924578-A2.
XX
XX 20-MAY-1999.
PD
XX
```

```

PF 09-OCT-1998; 98WO-IB01665.
XX
PR 01-SEP-1998; 98GB-0019016.
PR 06-NOV-1997; 97GB-0023316.
PR 14-NOV-1997; 97GB-0024190.
PR 18-NOV-1997; 97GB-0024386.
PR 27-NOV-1997; 97GB-0025158.
PR 10-DEC-1997; 97GB-0026147.
PR 14-JAN-1998; 98GB-0000759.
XX
XX (CHIR-) CHIRON SPA.
XX
XX Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX
DR WPI; 1999-327407/27.
DR N-PSDB; Z12253.
XX
XX Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
XX diagnosis, treatment and prevention of infection
XX
PS Claim 4; Page 371; 524pp; English.
XX
CC Amino acid sequences Y38499-Y38944 represent Neisseria meningitidis
CC and N. gonorrhoeae antigenic proteins. They are encoded by open
CC reading frames (ORFs) Z11972-Z12358. The antigenic proteins,
CC their fragments, their nucleic acids and antibodies are used for
CC diagnosis, prevention (as vaccines) or treatment of Neisseria
CC infections, such as meningitis, septicemia and gonorrhea. Both
CC organisms are closely related. Fragments of the nucleic acids
CC are useful as hybridisation probes and antisense reagents.
XX
SQ Sequence 1468 AA;

Query Match 22.9%; Score 128.5; DB 20; Length 1468;
Best Local Similarity 35.4%; Pred. No. 6.1e-06;
Matches 28; Conservative 15; Mismatches 27; Indels 9; Gaps 1;

QY 26 VTENNNDWVFMTQEEAKKNAMNHN--NORISGFSGFCEENGKNGALNLFNFKSAON 85
   : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 569 itgnkdltttgnn-----nmldskelayngwfgexdatktnglnlnyppeadr 619
   : : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 86 RFLTGGTNLNGKISVTQG 104
   | | | | | | | | : |
Db 620 tlllsggtlnngnitqng 638

RESULT 10
R92768
ID R92768 standard; Protein; 1394 AA.
XX
AC R92768;
XX
XX 01-JUN-1996 (first entry)
DT
DE Adhesion and penetration protein.
XX
XX Adhesion and penetration protein; hap gene; signal peptide;
KW protease; helper domain; pore; outer membrane protein; beta-domain;
KW secretion; active site; catalytic domain; recombinant vaccine;
KW monoclonal antibody; diagnostic; immunoassay.
XX
XX Haemophilus influenzae.
XX
XX Key Location/Qualifiers
FH Peptide 1..25
   /note= "Signal peptide"
FT
FT Peptide 27..33
   /note= "Mature protein N-terminal peptide"
FT
FT Protein 26..959
   /note= "Secreted 110-kDa protease fragment"
FT
FT Domain 241..248
   /note= "Putative catalytic domain"
FT
```

```

FT Active-site 243 /note= "Active site serine residue"
FT Region 698 /note= "Conserved cysteine residue"
FT Region 709 /note= "Conserved cysteine residue"
FT Cleavage-site 960 /note= "Site for cleavage of secreted protease from
FT outer membrane protein"
FT Protein 960..1394
FT /note= "45-kDa outer membrane protein fragment"
XX
XX W09605858-A1.
XX
XX 29-FEB-1996.
XX
XX 16-AUG-1995; 95WO-US10661.
XX
XX 25-AUG-1994; 94US-0296791.
XX
XX (STRD ) UNIV LELAND STANFORD JUNIOR.
XX (UNIW ) UNIV WASHINGTON.
XX Falkow S, St Geme JW;
XX WPI; 1996-151147/15.
XX N-PSDB; T17215.
XX
XX Haemophilus adhesion and penetration protein and corresponding DNA
PT - used to produce vaccines against H. influenzae infection
XX
XX Claim 2; Fig 6; 105pp; English.
XX
XX The sequence represents a Haemophilus influenzae adhesion and
XX penetration protein, encoded by the hap gene. The protein is
XX first synthesised as a preprotein with 3 functional domains (the
XX N-terminal signal peptide, the protease and a C-terminal helper
XX domain), which is transported to the periplasm, followed by
XX insertion of the C-terminal beta-domain into the outer membrane,
XX possibly forming a pore, and export of the N-terminal fragment
XX through the outer membrane, followed by autoproteolytic cleavage
XX and secretion of the mature protease, leaving an outer membrane
XX protein fragment. An active site serine residue characteristic of
XX serine proteases is present, along with 2 cysteine residues which
XX may be important in forming the catalytic domain. A gene encoding
XX the protein may be inserted in a vector and expressed in
XX recombinant host cells, for use as a recombinant vaccine. The
XX protein may also be used in diagnostic monoclonal antibody production.
XX
XX Sequence 1394 AA;

```

```

Query Match 20.3%; Score 114; DB 17; Length 1394;
Best Local Similarity 38.3%; Pred. No. 0.00027;
Matches 23; Conservative 13; Mismatches 24; Indels 0; Gaps 0;

```

```

Qy 45 NAMNHKNQIRISGFGFGENGKGNALNFKGSAQNRFLLTGTTNLNGKISVTQG 104
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 551 nninkldyrkeiayngwfgetdknkhgrlnlykpttedrtlllsgntlnkgdtkg 610

```

```

RESULT 11
Y38824
ID Y38824 standard; Protein; 1449 AA.
XX
XX Y38824;
AC
XX 08-OCT-1999 (first entry)
XX
XX Neisseria meningitidis strain A antigen encoded by ORF1.
DE
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW
KW treatment; Neisseria infection; meningitis; septicaemia; gonorrhea.

```

```

XX Neisseria meningitidis.
OS
XX W09924578-A2.
PN
XX 20-MAY-1999.
PD
XX
XX 09-OCT-1998; 98WO-IB01665.
PF
XX
XX 01-SEP-1998; 98GB-0019016.
PR
XX 06-NOV-1997; 97GB-0023516.
PR
XX 14-NOV-1997; 97GB-0024190.
PR
XX 18-NOV-1997; 97GB-0024386.
PR
XX 27-NOV-1997; 97GB-0025158.
PR
XX 10-DEC-1997; 97GB-0026147.
PR
XX 14-JAN-1998; 98GB-0000759.
XX
XX (CHIR-) CHIRON SPA.
FA
XX
XX Grandi G, Masiagnani V, Pizsa M, Rappuoli R., Scariato V;
PI
XX
XX WPI; 1999-327407/27.
DR
XX N-PSDB; Z12252.
DR
XX
XX Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT diagnosis, treatment and prevention of infection
PT
XX
XX Claim 4; Page 365; 524pp; English.
XX
XX Amino acid sequences Y38499-Y38944 represent Neisseria meningitidis
CC and N. gonorrhoeae antigenic proteins. They are encoded by open
CC reading frames (ORFs) Z11972-Z12358. The antigenic proteins,
CC their fragments, their nucleic acids and antibodies are used for
CC diagnosis, prevention (as vaccines) or treatment of Neisseria
CC infections, such as meningitis, septicaemia and gonorrhea. Both
CC organisms are closely related. Fragments of the nucleic acids
CC are useful as hybridisation probes and antisense reagents.
XX
XX Sequence 1449 AA;
SQ

```

```

Query Match 20.3%; Score 114; DB 20; Length 1449;
Best Local Similarity 38.8%; Pred. No. 0.00028;
Matches 26; Conservative 11; Mismatches 28; Indels 2; Gaps 1;

```

```

Qy 38 TOEEAKKNAHKNQIRISGFGFGENGKGNALNFKGSAQNRFLLTGTTNLNG 97
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 567 tqpsgkn--inrlsyskeiayngwfgekdtktngrlnlyvqpaaedrxtllsgtning 624

```

```

Qy 98 KISVTQG 104
| : | |
Db 625 nitqtng 631

```

```

RESULT 12
Y56621
ID Y56621 standard; Protein; 1454 AA.
XX
XX AC Y56621;
AC
XX
XX 21-FEB-2000 (first entry)
DT
XX
XX Neisseria meningitidis strain ATCC13090 BASB006 protein sequence.
DE
XX
XX Neisseria meningitidis; BASB006; diagnosis; bacterial; infection;
KW vaccine; antibiotic; upper respiratory tract infection; meningitis;
KW invasive bacterial disease; bacteraemia; screening; antibacterial.
XX
XX Neisseria meningitidis.
OS
XX
XX W09955873-A2.
PN
XX
XX 04-NOV-1999.
PD

```

PD	20-MAY-1999.
XX	
PF	09-OCT-1998; 98WO-IB01665.
XX	
PR	01-SEP-1998; 98GB-0019016.
PR	06-NOV-1997; 97GB-0023516.
PR	14-NOV-1997; 97GB-0024190.
PR	18-NOV-1997; 97GB-0024386.
PR	27-NOV-1997; 97GB-0025158.
PR	10-DEC-1997; 97GB-0026147.
PR	14-JAN-1998; 98GB-0000759.
XX	
PA	(CHIR-) CHIRON SPA.
XX	
PI	Grandi G, Masignani V, Pizza M, Rappuoli R, Scarlato V;
XX	
DR	WPI: 1999-327407/27.
XX	N-PSDB; Z12251.
DR	
XX	
P	Proteins from Neisseria meningitidis and N. gonorrhoeae useful for
PT	diagnosis, treatment and prevention of infection
XX	
PS	Claim 4; Page 361; 524pp; English.
XX	
CC	Amino acid sequences Y38499-Y38944 represent Neisseria meningitidis
CC	and N. gonorrhoeae antigenic proteins. They are encoded by open
CC	reading frames (ORFs) Z11972-Z12358. The antigenic proteins,
CC	their fragments, their nucleic acids and antibodies are used for
CC	diagnosis, prevention (as vaccines) or treatment of Neisseria
CC	infections, such as meningitis, septicaemia and gonorrhea. Both
CC	organisms are closely related. Fragments of the nucleic acids
CC	are useful as hybridisation probes and antisense reagents.
XX	
SQ	Sequence 1457 AA;
	Query Match 19.8%; Score 111.5; DB 20; Length 1457;
	Best Local Similarity 34.2%; Pred. No. 0.00056;
	Matches 27; Conservative 13; Mismatches 30; Indels 9; Gaps
OY	26 VTENNDFVMGYTOEPAKKNNHKNRISGFSFGFGENCKGHNGALNLFNGKSAON 85

569	ifgnkalatg-----nnnsldskke-----layngwigeakdtktngnrlnlvyqpaaedr	611
QY	86	RELLTGGTNLNGKISVTQG 104
Db	620	tllsggtlnqnitqng 638
RESULT	14	
B58592		
ID	B58592	standard; Protein; 1457 AA.
XX	AC	B58592;
XX	AC	
DT	DT	
DT	13-MAR-2001	(first entry)
DE	N.	meningitidis amino acid sequence orfl-1.pep SEQ ID NO:87.
XX		
KW	Neisseria	meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KW	diagnosis; antigen; detection; infection; gene therapy; antibacterial	
XX	OS	
OS	Neisseria	meningitidis.
XX		
PN	WO2000066791-A1.	
XX		
PPD	09-NOV-2000.	
XX		
PF	08-MAR-2000; 2000WO-US05928.	
XX		
PR	30-APR-1999; 99US-0132068.	
PR	08-OCT-1999; 99WO-US23573.	
PR	28-FEB-2000; 2000GB-0004695.	

```

XX (CHIR ) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;
XX WPI; 2000-647603/62.
DR N-PSDB; F21587.
XX
PT Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections -
XX
PS Example 1; Page 114; 692pp; English.
XX
CC The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in F21544 and F21607 to
CC F21613 represent fragments of the NMB genomic sequence, as the sequence
CC was too long to go in a record on its own it was split into 8 sequences
CC which overlap each other at the beginning and end of each sequence by
CC 4980 bp (i.e. the last 4980 bp of F21544 is repeated at the beginning
CC of F21607, the last 4980 bp of F21607 are repeated at the beginning of
CC F21608, and so on). F21545 to F21588 encode the Neisseria proteins given
CC in B58550 to B58593, and F21589 to F21606 represent PCR primers which
CC are used in the exemplification of the present invention. The NMB genome
CC and fragments from it have antibacterial activity, and can be used in
CC vaccines and gene therapy. Neisseria nucleic acids, proteins and/or
CC antibodies which binds to the proteins can be used in compositions for
CC treating or preventing infection due to Neisserial bacteria or as a
CC diagnostic reagent for detecting the presence of Neisserial bacteria or
CC of antibodies raised to Neisserial bacteria. Computers, computer memory,
CC computer storage medium or computer databases can be used in a search to
CC identify open reading frames (ORFs) or coding sequences within the NMB
CC genome. The DNA sequences provide further opportunities to find
CC antigenic or immunogenic proteins which are more effective in vaccines
XX
SQ Sequence 1457 AA;

Query Match 19.8%; Score 111.5; DB 21; Length 1457;
Best Local Similarity 34.2%; Pred. No. 0.00056;
Matches 27; Conservative 13; Mismatches 30; Indels 9; Gaps 2;

QY 26 VTENNDFVFMGYTQEEAKKNAMNHNQRIISGFSGFFGEGNGKGNALNLFNGKSAQN 85
Db 569 itgnkdiattg-----nnnsidskke---layngwfgekdtktngnlrvyqpaaedr 619

QY 86 RFLLTGGTNLNGKISVTQG 104
Db 620 tlllsggtlnlgnitqtng 638

RESULT 15
ID B25662 standard; Protein; 1457 AA.
XX
AC B25662;
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis amino acid sequence orfi-1.pep SEQ ID NO:1047.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX

```

08-OCT-1999; 99WO-US23573.

09-OCT-1998; 98US-0103794.

30-APR-1999; 99US-0132068.

(CHIR) CHIRON CORP.

Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;

Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;

Rappuoli R, Pizza M;

WPI; 2000-318079/27.

N-PSDB; A81302.

Isolated nucleotide sequences of Neisseria meningitidis which can be used in the diagnosis and treatment of N. meningitidis infection and other Neisserial infections, for example, N.gonorrhoea -

Example 1; Page 113; 1760pp; English.

The present invention describes methods of obtaining immunogenic proteins from Neisseria genomic sequences. A81453 to A82414 represent specifically claimed Neisseria meningitidis genomic DNA sequences; A81260 to A81303 and B25620 to B25663 represent Neisseria DNA sequences and their corresponding proteins; A81254 to A81259 and A81304 to A81321 represent PCR primers used in the isolation of Neisseria meningitidis DNA sequences; and A81322 to A81452 represent Neisseria meningitidis MenB polynucleotide ORF sequences, which are all used in the exemplification of the present invention. The nucleic acid sequences, protein sequences, of the antibodies against them, can be used in the manufacture of a composition. The composition can be used as a medicament (or in the manufacture of a medicament) for treating, preventing or diagnosing infection due to Neisserial bacteria. For example, some of the identified proteins could be components of vaccines against Meningococcus B; against all serotypes; and/or against all pathogenic Neisseriae. Identification of sequences from the bacterium will also facilitate production of biological probes, particularly organism-specific probes. Attempts to make efficacious Meningococcus B vaccines have failed mainly due to antigen tolerance. Multivalent vaccines have also been tried but none have successfully overcome antigenic variability. The provision of further, complete sequences may provide an opportunity to identify secreted or surface exposed proteins that may be presumed targets for the immune system and which are not antigenically variable or at least more conserved than other more variable regions.

Sequence 1457 AA;

Query Match 19.8%; Score 111.5; DB 21; Length 1457;

Best Local Similarity 34.2%; Pred. No. 0.00056;

Matches 27; Conservative 13; Mismatches 30; Indels 9; Gaps 2;

QY 26 VTENNDFVFMGYTQEEAKKNAMNHNQRIISGFSGFFGEGNGKGNALNLFNGKSAQN 85

Db 569 itgnkdiattg-----nnnsidskke---layngwfgekdtktngnlrvyqpaaedr 619

QY 86 RFLLTGGTNLNGKISVTQG 104

Db 620 tlllsggtlnlgnitqtng 638

Search completed: May 1, 2001, 14:58:48

Job time: 610 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 14:59:31 ; Search time 37.5 seconds
(without alignments)
53.278 Million cell updates/sec

Title: US-09-142-970-5
Perfect score: 565
Sequence: 1 LYKKNRYALKSGRLNAP.....NRFLLTGGANLNGKISVTQG 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 185757 seqs, 19210857 residues

Total number of hits satisfying chosen parameters: 185757

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/1aa/5A-COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B-COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A-COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B-COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCTUS-COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	556	98.4	1507	6	Patent No. 5268270
2	298.5	52.8	1541	5	Sequence 3, Appli
3	291.5	51.6	1545	5	Sequence 4, Appli
4	291.5	51.6	1702	5	Sequence 5, Appli
5	281.5	49.8	1848	5	Sequence 6, Appli
6	101	17.9	1394	5	Sequence 2, Appli
7	72.5	12.8	605	1	Sequence 3, Appli
8	72.5	12.8	605	1	Sequence 3, Appli
9	72.5	12.8	605	2	Sequence 3, Appli
10	70	12.4	1577	2	Sequence 3, Appli
11	67.5	11.9	631	1	Sequence 11, App
12	67.5	11.9	631	1	Sequence 11, App
13	67.5	11.9	631	2	Sequence 11, App
14	67.5	11.9	631	2	Sequence 11, App
15	67.5	11.9	631	3	Sequence 11, App
16	67.5	11.9	631	3	Sequence 11, App
17	66.5	11.8	631	3	Sequence 11, App
18	66.5	11.8	631	2	Sequence 11, App
19	66.5	11.8	631	2	Sequence 11, App
20	66.5	11.8	631	2	Sequence 11, App
21	66.5	11.8	631	3	Sequence 11, App
22	66.5	11.8	631	3	Sequence 11, App
23	66	11.7	1536	1	Sequence 11, App
24	66	11.7	1536	1	Sequence 11, App
25	66	11.7	1536	2	Sequence 11, App
26	66	11.7	1536	2	Sequence 11, App
27	66	11.7	1536	2	Sequence 11, App

28	66	11.7	1536	2	US-08-617-697-2	Sequence 2, Appli
29	64.5	11.4	644	1	US-08-487-890A-6	Sequence 6, Appli
30	64.5	11.4	644	2	US-08-478-435-6	Sequence 6, Appli
31	64.5	11.4	644	2	US-08-337-483-6	Sequence 6, Appli
32	64.5	11.4	644	2	US-08-478-373-6	Sequence 6, Appli
33	64.5	11.4	644	3	US-08-474-671-6	Sequence 6, Appli
34	64.5	11.4	644	3	US-08-483-577A-6	Sequence 6, Appli
35	64.5	11.4	671	2	US-08-737-716-13	Sequence 13, Appli
36	63.5	11.2	421	2	US-08-807-263-4	Sequence 4, Appli
37	63.5	11.2	532	2	US-08-899-324-33	Sequence 33, Appli
38	63.5	11.2	532	4	US-08-329-892B-33	Sequence 33, Appli
39	63.5	11.2	597	1	US-08-462-884A-1	Sequence 1, Appli
40	63.5	11.2	597	1	US-08-461-881B-1	Sequence 1, Appli
41	63.5	11.2	597	2	US-09-123-960-1	Sequence 1, Appli
42	63.5	11.2	759	2	US-08-450-351-2	Sequence 2, Appli
43	63.5	11.2	759	2	US-08-450-351-4	Sequence 2, Appli
44	63	11.2	1290	1	US-08-138-641-2	Sequence 2, Appli
45	63	11.2	1290	1	US-08-138-133-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
5268270-2
; Patent No. 5268270
; APPLICANT: Meyer, Thomas F.; Halter, Roman; Pohlner, Johannes
; TITLE OF INVENTION: PROCESS FOR PRODUCING PROTEINS USING GRAM
; NEGATIVE HOST CELLS
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/171,872
; FILING DATE: 01-JUL-1987
; SEQ ID NO:2:
; LENGTH: 1507
5268270-2

Query Match 98.4%; Score 556; DB 6; Length 1507;
Best Local Similarity 99.0%; Pred. No. 1e-58;
Matches 103; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LYKKNRYALKSGRLNAPMPENGVAENNDWIFMGYTQEEARKNHNHNRIGDGG 60
Db 559 LYKKNRYALKSGRLNAPMPENGVAENNDWIFMGYTQEEARKNHNHNRIGDGG 618
QY 61 PFDEENGKHGALNLFNGKSAQNRFLLTGGANLNGKISVTQG 104
Db 619 PFDEENGKHGALNLFNGKSAQNRFLLTGGANLNGKISVTQG 662

RESULT 2
PCT-US95-10661A-3
; Sequence 3, Application PC/TUS9510661A
; GENERAL INFORMATION:
; APPLICANT: Washington University, et al.
; TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patenting Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10661A
; FILING DATE: 16-AUG-1995

```
;
; CLASSIFICATION:
; PRIOR APPLICATION DATA: US 08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: FP-59941/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1541 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; PCT-US95-10661A-3

Query Match 52.8%; Score 298.5; DB 5; Length 1541;
Best Local Similarity 51.9%; Pred. No. 1.7e-27;
Matches 54; Conservative 18; Mismatches 31; Indels 1; Gaps 1;

QY 1 LYYKNRYVYALKSGGRLNAPMPENGVAENNDWIFMGYTOEARKNAMNHKNNRRIGDFGG 60
DB 595 LNLNYYTALPKGASTRSELPKNSGESNENWLYMGKTSDEAKRNMHNNRNGFNG 654

QY 61 FFDEENGKNGHGNALNLFNGKSAQNRFLTTGGANLNGKISVTQG 104
DB 655 YFGEDEGK-NGNLNVTFKKGSEQNRFLLTGGTNLNGDLTVKKG 697

RESULT 3
PCT-US95-10661A-4
; Sequence 4, Application PC/TUS9510661A
; GENERAL INFORMATION:
; APPLICANT: Washington University, et al.
; TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10661A
; FILING DATE: 16-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: FP-59941/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 25
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1545 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; PCT-US95-10661A-5
```

```
;
; TOPOLOGY: unknown
; PCT-US95-10661A-4

Query Match 51.6%; Score 291.5; DB 5; Length 1545;
Best Local Similarity 52.8%; Pred. No. 1.2e-26;
Matches 56; Conservative 16; Mismatches 31; Indels 3; Gaps 2;

QY 1 LYYKNRYVYALKSGGRLNAPMPENGVAENNDWIFMGYTOEARKNAMNHKNNRRIGDF 58
DB 597 LYPNEENRYTALKKADASIRSEFPQNRGESNNWLYMGTEKADAKQKAMNHNNRNGMF 656

QY 59 GGFEDEENGKNGHGNALNLFNGKSAQNRFLTTGGANLNGKISVTQG 104
DB 657 NGYFGEDEGK-NGNLNVTFKKGSEQNRFLLTGGTNLNGDLNVQOG 701

RESULT 4
PCT-US95-10661A-5
; Sequence 5, Application PC/TUS9510661A
; GENERAL INFORMATION:
; APPLICANT: Washington University, et al.
; TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10661A
; FILING DATE: 16-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: FP-59941/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 5
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1702 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; PCT-US95-10661A-5
```

```
Query Match 51.6%; Score 291.5; DB 5; Length 1702;
Best Local Similarity 51.0%; Pred. No. 1.4e-26;
Matches 53; Conservative 17; Mismatches 33; Indels 1; Gaps 1;

QY 1 LYYKNRYVYALKSGGRLNAPMPENGVAENNDWIFMGYTOEARKNAMNHKNNRRIGDFGG 60
DB 601 LNLNYYTALPKGASTRSELPKNSGESNENWLYMGKTSDEAKRNMHNNRNGFNG 660

QY 61 FFDEENGKNGHGNALNLFNGKSAQNRFLTTGGANLNGKISVTQG 104
DB 661 YFGEDEGK-NGNLNVTFKKGSEQNRFLLTGGTNLNGDLKLVKKG 703
```



```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10661A
; FILING DATE: 16-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: FP-59941/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-10661A-2

Query Match 17.9%; Score 101; DB 5; Length 1394;
Best Local Similarity 36.7%; Pred.No. 0.0013;
Matches 22; Conservative 11; Mismatches 27; Indels 0; Gaps 0

QY 45 NAMNKHNNRRIGDFGFFDEENGKGHNGALNLFNGKSAQNRELLTGGANLNGKISVTQG 104
Db 551 NNINKLDYRKIEIAYWFGSETDKNKHGRNLNIYKPTEDRTILLSSGTNLKGDITQTGK 610

RESULT 7
US-08-462-884A-3
; Sequence 3, Application US/08462884A
; Patent No. 5624836
; GENERAL INFORMATION:
; APPLICANT: Lange III, Louis G
; TITLE OF INVENTION: Spilburg, Curtis A
; TITLE OF INVENTION: Mamalian Pancreatic Cholesterol
; TITLE OF INVENTION: Esterase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Banner & Allegretti
; STREET: 10 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/462,884A
; FILING DATE: Unknown
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: A. Blair Hughes
; REGISTRATION NUMBER: 32,901
; REFERENCE/DOCKET NUMBER: 89,852-J
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/715-1000
; TELEFAX: 312/715-1234
; TELEX: 910/221-5317
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 605 amino acids

```

;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-462-884A-3

Query Match 12.8%; Score 72.5; DB 1; Length 605;
Best Local Similarity 24.7%; Pred. No. 1.2;
Matches 24; Conservative 16; Mismatches 40; Indels 17; Gaps 2;

Qy 17 LNAPMPENGVAENNDWIFMGYTQEEARKNMKNRRIGDGGFFDEENCKGHNGALNL 76
Db 91 LQATLTQDSTYGNEDCLYLNWVPGRKE-VSHDLPVMIWYIGGAFLMGSGGANFLKNY 149

Qy 77 NFNGKSAQNR-----FLLTGGANLNG 97
Db 150 LYDGEIATRGNVIVVTNRYVRGPGFLSTGDSNLP 186

RESULT 8
US-08-461-881B-3
; Sequence 3, Application US/08461881B
; Patent No. 5792832
; GENERAL INFORMATION:
; APPLICANT: Lange III, Louis G
; APPLICANT: Spilburg, Curtis A
; TITLE OF INVENTION: Mammalian Pancreatic
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" hard disc
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,881B
FILING DATE: June 5, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: A. Blair Hughes
REGISTRATION NUMBER: 32,901
REFERENCE/DOCKET NUMBER: 89,852-J
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/913-0001
TELEX: 910/221-5317
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 605 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-881B-3

Query Match 12.8%; Score 72.5; DB 1; Length 605;
Best Local Similarity 24.7%; Pred. No. 1.2;
Matches 24; Conservative 16; Mismatches 40; Indels 17; Gaps 2;

Qy 17 LNAPMPENGVAENNDWIFMGYTQEEARKNMKNRRIGDGGFFDEENCKGHNGALNL 76
Db 91 LQATLTQDSTYGNEDCLYLNWVPGRKE-VSHDLPVMIWYIGGAFLMGSGGANFLKNY 149

Qy 77 NFNGKSAQNR-----FLLTGGANLNG 97
Db 150 LYDGEIATRGNVIVVTNRYVRGPGFLSTGDSNLP 186

RESULT 9
US-09-123-960-3
; Sequence 3, Application US/09123960
; Patent No. 5981299
; GENERAL INFORMATION:
; APPLICANT: Lange III, Louis G
; APPLICANT: Spilburg, Curtis A
; TITLE OF INVENTION: Mammalian Pancreatic
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" hard disc
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/123,960
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/461,881
FILING DATE: June 5, 1995
ATTORNEY/AGENT INFORMATION:
NAME: A. Blair Hughes
REGISTRATION NUMBER: 32,901
REFERENCE/DOCKET NUMBER: 89,852-K
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/913-0001
TELEFAX: 312/913-0002
TELEX: 910/221-5317
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 605 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-123-960-3

Query Match 12.8%; Score 72.5; DB 2; Length 605;
Best Local Similarity 24.7%; Pred. No. 1.2;
Matches 24; Conservative 16; Mismatches 40; Indels 17; Gaps 2;

Qy 17 LNAPMPENGVAENNDWIFMGYTQEEARKNMKNRRIGDGGFFDEENCKGHNGALNL 76
Db 91 LQATLTQDSTYGNEDCLYLNWVPGRKE-VSHDLPVMIWYIGGAFLMGSGGANFLKNY 149

Qy 77 NFNGKSAQNR-----FLLTGGANLNG 97
Db 150 LYDGEIATRGNVIVVTNRYVRGPGFLSTGDSNLP 186

RESULT 10
US-08-793-824-2
; Sequence 2, Application US/08793824
; Patent No. 5981838
; GENERAL INFORMATION:
; APPLICANT: Simpson, Christine Lynn
; APPLICANT: Giffard, Philip Morrison
; APPLICANT: Jacques, Nicholas Anthony
; TITLE OF INVENTION: Genetic Manipulation of Plants to
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Griffith Hack & Co
; STREET: Level 8, 160 Walker Street

Query Match 11.9%; Score 67.5; DB 1; Length 631;
Best Local Similarity 24.2%; Pred. No. 5.1;
Matches 29; Conservative 10; Mismatches 38; Indels 43; Gaps 5;

```

Query Match      11.9%; Score 67.5; DB 1; Length 631;
Best Local Similarity 24.2%; Pred. No. 5.1;
Matches 29; Conservative 10; Mismatches 38; Indels 43; Gaps

Qy 3 YKNVRYALKSGRLNAPMPENGVAE--NNDWIIFMGYTQEARKNAMNHKNNRRIGDPGGF 61
   ||| | : : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 150 YLGYYGVAFYVGNKTATNLPVSGIAKYKGTWDFITAT-----KNGQVSLFGSA 198

Qy 62 F-----DEENCKGHNGALNINFNCKSAQNRLITG-----GANLN 96
   ||| | : : : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 199 FGAVNRRSAISEDINLENNLKNAGAGITSEFTVNFQPK-----LTGKIYNERFTNLN 252

```

```

RESULT 12
US-08-478-435-111
; Sequence 111, Application US/08478435
; Patent No. 5922323
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena
; APPLICANT: Harkness, Robin
; APPLICANT: Schryvers, Anthony
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Yang, Yan-ping
; APPLICANT: Murdin, Andrew
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Transferrin Receptor Genes
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; City: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,435
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,483
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435

```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,116
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,968
; FILING DATE: 08-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-462 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 111:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 631 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-478-435-111
```

```
Query Match 11.9% Score 67.5; DB 2; Length 631;
Best Local Similarity 24.2%; Pred. No. 5.1;
Matches 29; Conservative 10; Mismatches 38; Indels 43; Gaps 5;

QY 3 YKNRYVALKSGGRLNAPMPENGVAE--NDWIFMGYTOEARKNAMHNRRIGDFGGF 61
| | | | | : | : | : | : | : | : | : | : | : | : | : |
Db 150 YLGYYGYAFYGNKTNLPSVGIKYGKGTWDFITAT-----KNGQYSLFGSA 198
| | | | | : | : | : | : | : | : | : | : | : | : | : |
QY 62 F-----DEENGKGHGALNLFNPKSAQNRLITG-----GANLN 96
| | | | | : | : | : | : | : | : | : | : | : | : | : |
Db 199 FGAYNRRSAISEDIDNLENNLKNAGLTSEFTVNFGTK-----LTGKLYYNERETNLN 252
```

```
RESULT 13
US-08-337-483-111
; Sequence 111, Application US/08337483
; Patent No. 5922562
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena
; APPLICANT: Harkness, Robin
; APPLICANT: Schryvers, Anthony
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Yang, Yan-ping
; APPLICANT: Murdin, Andrew
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Transferrin Receptor Genes
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/337,483
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-410 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
```

```
; INFORMATION FOR SEQ ID NO: 111:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 631 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-337-483-111

Query Match 11.9% Score 67.5; DB 2; Length 631;
Best Local Similarity 24.2%; Pred. No. 5.1;
Matches 29; Conservative 10; Mismatches 38; Indels 43; Gaps 5;

QY 3 YKNRYVALKSGGRLNAPMPENGVAE--NDWIFMGYTOEARKNAMHNRRIGDFGGF 61
| | | | | : | : | : | : | : | : | : | : | : | : | : |
Db 150 YLGYYGYAFYGNKTNLPSVGIKYGKGTWDFITAT-----KNGQYSLFGSA 198
| | | | | : | : | : | : | : | : | : | : | : | : | : |
QY 62 F-----DEENGKGHGALNLFNPKSAQNRLITG-----GANLN 96
| | | | | : | : | : | : | : | : | : | : | : | : | : |
Db 199 FGAYNRRSAISEDIDNLENNLKNAGLTSEFTVNFGTK-----LTGKLYYNERETNLN 252

RESULT 14
US-08-478-373-111
; Sequence 111, Application US/08478373
; Patent No. 5922841
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena
; APPLICANT: Harkness, Robin
; APPLICANT: Schryvers, Anthony
; APPLICANT: Chong, Pele
; APPLICANT: Gray-Owen, Scott
; APPLICANT: Yang, Yan-ping
; APPLICANT: Murdin, Andrew
; APPLICANT: Klein, Michel
; TITLE OF INVENTION: Transferrin Receptor Genes
; NUMBER OF SEQUENCES: 147
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/478,373
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/337,483
; FILING DATE: 08-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/175,116
; FILING DATE: 29-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/148,968
; FILING DATE: 08-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-463 MIS:vg
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 111:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 631 amino acids
```

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-478-373-111

Query Match 11.9%; Score 67.5; DB 2; Length 631;
Best Local Similarity 24.2%; Pred. No. 5.1;
Matches 29; Conservative 10; Mismatches 38; Indels 43; Gaps 5;
QY 3 YKNRYVALKSGGRLNAPMPENGVAE--NNDWIFMGYTOEEARKNMHKNRRIGDFFGGF 61
Db 150 YLGYYGYAFYGNKATNLPVSGIAKYGTWDFITAT-----KNGQYSLEFGSA 198
QY 62 F-----DEENGKGHGALNLFNFKSAQNRFLLTG-----GANLN 96
Db 199 FGAYNRRSAISEDIDNLENNLKNGAGLTSEFTVNFQTK-----LTGKLYYNERETNLN 252

RESULT 15
US-08-474-671-111
Sequence 111, Application US/08474671
Patent No. 6008326
GENERAL INFORMATION:
APPLICANT: Loosmore, Sheena
APPLICANT: Harkness, Robin
APPLICANT: Schryvers, Anthony
APPLICANT: Chong, Pele
APPLICANT: Gray-Owen, Scott
APPLICANT: Yang, Yan-Ping
APPLICANT: Murdin, Andrew
APPLICANT: Klein, Michel
TITLE OF INVENTION: Transferrin Receptor Genes
NUMBER OF SEQUENCES: 147
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: Suite 701, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,671
FILING DATE: 07-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/337,483
FILING DATE: 08-NOV-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/175,116
FILING DATE: 29-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/148,968
FILING DATE: 08-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-465 MIS:vg
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 111:
SEQUENCE CHARACTERISTICS:
LENGTH: 631 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-474-671-111

Query Match 11.9%; Score 67.5; DB 3; Length 631;
Best Local Similarity 24.2%; Pred. No. 5.1;
Matches 29; Conservative 10; Mismatches 38; Indels 43; Gaps 5;
QY 3 YKNRYVALKSGGRLNAPMPENGVAE--NNDWIFMGYTOEEARKNMHKNRRIGDFFGGF 61
Db 150 YLGYYGYAFYGNKATNLPVSGIAKYGTWDFITAT-----KNGQYSLEFGSA 198
QY 62 F-----DEENGKGHGALNLFNFKSAQNRFLLTG-----GANLN 96
Db 199 FGAYNRRSAISEDIDNLENNLKNGAGLTSEFTVNFQTK-----LTGKLYYNERETNLN 252

Search completed: May 1, 2001, 14:59:32
Job time: 639 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 15:03:34 ; Search time 26.69 seconds
(without alignments)
133.480 Million cell updates/sec

Title: US-09-142-970-2

Perfect score: 562

Sequence: 1 LYKRYRYALKSGSVNAP.....NRFLTGCTNLNGKISVTQG 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	508	90.4	1532	1	IGA_NEIGO
2	329.5	58.6	1541	1	IGAL_HAEIN
3	328.5	58.5	1694	1	IGA0_HAEIN
4	328.5	58.5	1702	1	IGA2_HAEIN
5	319.5	56.9	1545	1	IGA3_HAEIN
6	318.5	56.7	1849	1	IGA4_HAEIN
7	115.5	20.6	1409	1	HAP1_HAEIN
8	114	20.3	1394	1	HAP2_HAEIN
9	91.5	16.3	989	1	PTP3_DICDI
10	83.5	14.9	451	1	ARP2_PLAFA
11	77	13.7	625	1	TBP2_HAEIN
12	77	13.7	712	1	TBPB_NEIMB
13	77	13.7	982	1	CAPP_ANASP
14	74.5	13.3	719	1	ARP_YEAST
15	74.5	13.3	1029	1	KSP1_YEAST
16	74.5	13.3	1596	1	MAM_DROME
17	74	13.2	1243	1	VG37_BPK3
18	73.5	13.1	671	1	ALYS_ENTFA
19	71	12.6	824	1	ROU_HUMAN
20	70.5	12.5	700	1	NONA_DROME
21	70.5	12.5	882	1	IF2_BORBU
22	70	12.5	2334	1	WAPA_BACSU
23	69	12.3	215	1	RL4_HELPJ
24	69	12.3	215	1	RL4_HELPY
25	69	12.3	563	1	SRGB_CHICK
26	68.5	12.2	338	1	G3P_PHARH
27	68.5	12.2	954	1	XYNA_RUMFL
28	68.5	12.2	2339	1	RPC1_PLAFA
29	68	12.1	537	1	ARP_PLAFA
30	68	12.1	1444	1	ADP1_MYCGE
31	67	11.9	550	1	SRGB_HUMAN
32	67	11.9	826	1	YN83_CAEEL
33	66.5	11.8	101	1	VPR_STIVSP

RESULT 1

ID	IGA_NEIGO	STANDARD	PRT	1532 AA
AC	P09790			
DT	01-MAR-1989	(Rel. 10, Created)		
DT	01-MAR-1989	(Rel. 10, Last sequence update)		
DT	01-OCT-2000	(Rel. 40, Last annotation update)		
DE	IGA-SPECIFIC SERINE ENDOPEPTIDASE PRECURSOR (EC 3.4.21.72) (IGA PROTEASE).			
DE	IGA.			
GN	Neisseria gonorrhoeae.			
OS	Neisseria gonorrhoeae.			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_Taxid=485;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	STRAIN=MS11;			
RX	MEDLINE=87115823; PubMed=3027577;			
RA	Pohlner J., Halter R., Beyreuther K., Meyer T.F.;			
RT	"Gene structure and extracellular secretion of Neisseria gonorrhoeae Iga protease.";			
RL	Nature 325:458-462(1987).			
RN	[2]			
RP	ACTIVE SITE.			
RX	MEDLINE=90154052; PubMed=2105953;			
RA	Bachovchin W.W., Plaut A.G., Flentke G.R., Lynch M., Kettner C.A.;			
RT	"Inhibition of Igal proteinases from Neisseria gonorrhoeae and Hemophilus influenzae by peptide prolyl boronic acids.";			
RL	J. Biol. Chem. 265:3738-3743(1990).			
CC	-!- FUNCTION: THIS PROTEASE IS SPECIFIC FOR IMMUNOGLOBULIN A.			
CC	-!- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE SUBSTRATES ARE KNOWN.			
CC	-!- SUBCELLULAR LOCATION: SECRETED.			
CC	-!- DOMAIN: THE SIGNAL PEPTIDE GUIDE THE PRECURSOR TO THE PERIPLASMIC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS.			
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	EMBL; X04835; CAA28538.1; ..			
DR	PIR; A26039; A26039.			
DR	MEROPS; S06.001; ..			
DR	InterPro; IPR000710; ..			
DR	PRINTS; PR00921; IGASERPTASE.			
KW	Hydrolase; Serine protease; Zymogen; Autocatalytic cleavage;			
KW	Transmembrane; Signal.			
FT	SIGNAL 1 27			
FT	CHAIN 28 986 IGA-SPECIFIC SERINE ENDOPEPTIDASE.			

P34291 caenorhabdi
Q02480 human papil
P41073 drosophila
Q00662 dianthus ca
O83047 treponema p
P592m8 drosophila
Q26600 schistosoma
Q26600 schistosoma
P20148 neisseria g
P77581 escherichia
P08779 homo sapien
P04635 staphylococ

ALIGNMENTS

34	66.5	11.8	401	1	YK03_CAEEL
35	66.5	11.8	507	1	VL1_HPVO9
36	66.5	11.8	716	1	PEP_DROME
37	66.5	11.8	731	1	BGAL_DIACA
38	66	11.7	464	1	DNAA_TREPA
39	66	11.7	1080	1	HDC_DROME
40	66	11.7	1556	1	GLTS_SYNY3
41	65.5	11.7	280	1	SMX1_SCHMA
42	65.5	11.7	348	1	OMB2_NEIGO
43	65.5	11.7	406	1	ARGM_ECOLI
44	65.5	11.7	469	1	K1CX_HUMAN
45	65.5	11.7	641	1	LIP_STAHY

!- SMILLAKIII: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).

CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).

CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

-----NOT_ANNOTATED_CDS-----

DR EMBL; U32710; ?; NOT_ANNOTATED_CDS.
DR TIGR; H10248; ?
DR InterPro; IPR000710; ?
DR PRINTS; PR00921; IGASERPTASE.
KW Hydrolyase; Serine protease; Transmembrane; Zymogen; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 ? ADHESION AND PENETRATION PROTEIN.
FT PROPEP 2 1409 HELPER PEPTIDE (POTENTIAL).
FT ACT_SITE 250 250 BY SIMILARITY.
SQ SEQUENCE 1409 AA; 156797 MW; 63ABC893FA84D16E CRC64;

Query Match 20.6%; Score 115.5; DB 1; Length 1409;
Best Local Similarity 38.8%; Pred. No. 0.00074;
Matches 26; Conservative 12; Mismatches 28; Indels 1; Gaps 1;

QY 38 TOEAKKNNMHNKQIRISGFGFGKNGKNGHGNALNFGKSAQNRLLTGGTNLNG 97

Db 562 TAPSNKKN-INKLDSREIAYNGFGETDKNKHGRLNLIYKPTEDRTLLSGGTNLKG 620

QY 98 KISVTQ 104

Db 621 DITQTKG 627

RESULT 8

HAP_HAEIN STANDARD; PRT; 1394 AA.
ID HAP_HAEIN
AC P45387;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE ADHESION AND PENETRATION PROTEIN PRECURSOR (EC 3.4.21.-).
GN HAP.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus
OX NCBI_Taxid=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NHTI N187;
RX MEDLINE=95131744; PubMed=7830568;
RA St Geme J.W. III, de la Morena M.L., Falkow S.;
RT "A Haemophilus influenzae IGA protease-like protein promotes intimate interaction with human epithelial cells."
RL Mol. Microbiol. 14:217-233(1994).
CC -1- FUNCTION: PROBABLE PROTEASE; PROMOTES ADHERENCE AND INVASION BY DIRECTLY BINDING TO A HOST CELL STRUCTURE.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- DOMAIN: THE SIGNAL PEPTIDE GUIDE THE PRECURSOR TO THE PERIPLASMIC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; U11024; AAB03707.1; ?

DR InterPro; IPR000710; ?

DR PRINTS; PR00921; IGASERPTASE.

KW Hydrolyase; Serine protease; Transmembrane; Zymogen; Signal.

FT SIGNAL 1 25 POTENTIAL.

FT CHAIN 26 ? ADHESION AND PENETRATION PROTEIN.
FT PROPEP ? 1394 HELPER PEPTIDE (POTENTIAL).
FT ACT_SITE 243 243 BY SIMILARITY.
FT CONFLICT 1167 1167 MISSING (IN AAB03707).
SQ SEQUENCE 1394 AA; 155441 MW; 58F28660103F60F9 CRC64;

Query Match 20.3%; Score 114; DB 1; Length 1394;
Best Local Similarity 38.3%; Pred. No. 0.001;
Matches 23; Conservative 13; Mismatches 24; Indels 0; Gaps 0;

QY 45 NAMNHNKQIRISGFGFGKNGKNGHGNALNFGKSAQNRLLTGGTNLNGKISVTQ 104

Db 551 NNINKLDYRKEIAYNGFGETDKNKHGRLNLIYKPTEDRTLLSGGTNLKGDITQTKG 610

RESULT 9

PTP3_DICDI STANDARD; PRT; 989 AA.
ID PTP3_DICDI
AC P54637;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1997 (Rel. 35, Last annotation update)
DE PROTEIN-TYROSINE PHOSPHATASE 3 (EC 3.1.3.48) (PROTEIN-TYROSINE-DE PHOSPHATE PHOSPHOHYDROLASE 3).
GN (PTPC1 OR PTP3) AND (PTPC2 OR PTP3).
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_Taxid=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=AX3;
RX MEDLINE=96189126; PubMed=8628311;
RA Gamber M., Howard P.K., Hunter T., Firtel R.A.;
RT "Multiple roles of the novel protein tyrosine phosphatase PTP3 during Dictyostellium growth and development."
RL Mol. Cell. Biol. 16:2431-2444(1996).
CC -1- FUNCTION: SEEMS TO DEPHOSPHORYLATE A PROTEIN OF 130 KDA (P130).
CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
CC PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: IN THE ANTERIOR-LIKE AND PRESTALK CELL TYPES.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT MODERATE LEVELS DURING GROWTH AND DEVELOPMENT.
CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-TYROSINE PHOSPHATASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; U38197; AAC47041.1; ?

DR HSP; P18031; lpty.

DR DictyDb; DD01111; ptpc1.

DR DictyDb; DD07777; ptpc2.

DR InterPro; IPR000387; ?

DR InterPro; IPR000387; ?

DR Pfam; PF00102; Y_phosphatase; 1.

DR PRINTS; PR00700; PRTVPHPTASE.

DR PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.

DR PROSITE; PS00556; TYR_PHOSPHATASE_2; 1.

DR PROSITE; PS00555; TYR_PHOSPHATASE_PTP; 1.

KW Hydrolase.

FT ACT_SITE 649 649 BY SIMILARITY.

FT DOMAIN 64 71 POLY-ASN.

FT DOMAIN 109 118 POLY-ASN.

FT DOMAIN 137 190 POLY-ASN.

FT DOMAIN 249 257 POLY-SER.

FT DOMAIN 258 265 POLY-THR.


```
RESULT 12
TBPB_NEIMB
ID TBPB_NEIMB STANDARD; PRT; 712 AA.
AC QKQVO;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE TRANSFERRIN-BINDING PROTEIN 2 PRECURSOR (TBP-2).
GN TBPB OR TBP2 OR NMB0460.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; Pubmed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,
RA Cotton M.D., Uterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gili J., Scariato V., Massignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58."
RL Science 287:1809-1815(2000).
CC -1- FUNCTION: ACTS AS A TRANSFERRIN RECEPTOR AND IS REQUIRED FOR
CC TRANSFERRIN UTILIZATION (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID
CC ANCHOR (PROBABLE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE002402; AAP40897.1;
CC TIGR; NMB0460;
CC DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
CC KW Outer membrane; Receptor; Signal; Lipoprotein.
CC FT SIGNAL 1 20 BY SIMILARITY.
CC FT CHAIN 21 712 TRANSFERRIN-BINDING PROTEIN 2.
CC FT LIPID 21 21 N-ACYL DIGLYCERIDE (PROBABLE).
CC FT SEQUENCE 712 AA; 77416 MW; 77EA248941B8EFOC CRC64;
CC -----
Query Match 13.7%; Score 77; DB 1; Length 712;
Best Local Similarity 19.7%; Pred. No. 2.5;
Matches 29; Conservative 18; Mismatches 54; Indels 46; Gaps 4;
QY 3 YKNYRY-----YALKSGGSYNAPMPENG-VTERNDW 32
Db 149 YENFYVSGWFKYHAKREFNLKVPKSAKNGDDGYIFHGCKEPSRQLPASGKITRYKGVW 208
QY 33 VFMYGTQEEAKKNAM---NHKNORISGFSFGFFCEE-----NGKGHGALNL 76
Db 209 HFATDTKQKQFREIIQPSKSGDRYSFGSGDDGEEYSNKNKSTLTDCQEGYGFSTNLEV 268
QY 77 NFNGKSAQNRFLLTGGTNLNGKISVQTQ 103
Db 269 DFHNKLTGKLIRNANTDNNQATTQ 295
RESULT 13
CAPP_ANASP
ID CAPP_ANASP STANDARD; PRT; 982 AA.
AC P28594;
DT 01-DEC-1992 (Rel. 24, Created)
```

```
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PHOSPHOENOLPYRUVATE CARBOXYLASE (PEPCASE) (PEPC).
GN PPC.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92268848; Pubmed=1588304;
RA Luinburg I., Coleman J.R.;
RT Identification, characterization and sequence analysis of the gene
RT encoding phosphoenolpyruvate carboxylase in Anabaena sp. PCC 7120."
RL J. Gen. Microbiol. 138:685-691(1992).
CC -1- FUNCTION: TO FORM OXALOACETATE, A FOUR-CARBON DICARBOXYLIC ACID
CC SOURCE FOR THE TRICARBOXYLIC ACID CYCLE.
CC -1- CATALYTIC ACTIVITY: ORTHOPHOSPHATE + OXALOACETATE -> H(2)O +
CC PHOSPHOENOLPYRUVATE + CO(2).
CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.
CC -1- SIMILARITY: BELONGS TO THE PEPCASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M80541; AAA22023.1;
CC PIR; A44831; A44831.
CC DR InterPro; IPR001449;
CC DR Pfam; PF00311; PEPCase; 3.
CC DR PRINTS; PR00150; PEPCARBLASE.
CC DR PROSITE; PS00393; PEPCASE_2; 1.
CC DR PROSITE; PS00781; PEPCASE_1; 1.
CC KW Lyase; Carbon dioxide fixation; Allosteric enzyme;
CC KW Tricarboxylic acid cycle.
CC FT ACT_SITE 155 155 BY SIMILARITY.
CC FT ACT_SITE 628 628 BY SIMILARITY.
CC FT SEQUENCE 982 AA; 112679 MW; 250621FFAFC8DE CRC64;
CC -----
Query Match 13.7%; Score 77; DB 1; Length 982;
Best Local Similarity 21.6%; Pred. No. 3.6;
Matches 25; Conservative 20; Mismatches 51; Indels 20; Gaps 2;
QY 8 YALKSGGSVN-----APMPENGVTENNNDWVFMYGTQEEAKKNAMN---- 48
Db 579 YRALLAGGYKNTVEKVPNTLTQSPAPSPQSVLTPLDQLQVNLGYSDSNKDSGLSSNWE 638
QY 49 -HKNORISGFSFGFGENGKGHNGALNPNKSAQNRFLLTGGTNLNGKISVQTQ 103
Db 639 IHKAQKSLQIAEYVNLRIFFHGRGSGVGRGGGPAHEAILAQPGHSINGRIKITE 694
RESULT 14
ARP_YEAST
ID ARP_YEAST STANDARD; PRT; 719 AA.
AC P32770; Q12228;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE ARP PROTEIN.
GN ARP1 OR ARP OR YDL167C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AH22;
RX MEDLINE=93247548; Pubmed=8483449;
```

RA Wehner E.P., Rao E., Brendel M.;
RT "Molecular structure and genetic regulation of SFA, a gene
RT responsible for resistance to formaldehyde in Saccharomyces
RT cerevisiae, and characterization of its protein product.";
RL Mol. Gen. Genet. 237:351-358(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RA Pohl T.M.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 2 RANBP-TYPE ZINC FINGERS
CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RNM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X68020; CAA48159.1; -
DR EMBL: 267750; CAA91579.1; -
DR EMBL: 274215; CAA98741.1; -
DR PIR: S31139; S31139.
DR HSSP: P04170; 6RXN.
DR SGD: S0002326; NRX1.
DR InterPro: IPR000504; -
DR InterPro: IPR001876; -
DR Pfam: PF00076; rim; 1.
DR Pfam: PF00641; zf-RanBP; 2.
DR PROSITE: PS0102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP.1; FALSE_NEG.
KW Nuclear protein; zinc-finger; RNA-binding; Repeat.
FT DOMAIN 226 322 RNA-BINDING (RRM).
FT DOMAIN 490 564 ASN-RICH.
FT CONFLICT 493 493 I -> N (IN REF. 1).
SQ SEQUENCE 719 AA; 79299 MW; ADA9BC09FD582669 CRC64;

Query Match 13.3%; Score 74.5; DB 1; Length 719;
Best Local Similarity 21.9%; Pred. No. 4.5;
Matches 25; Conservative 9; Mismatches 35; Indels 45; Gaps 3;

QY 5 NYRYVALKSGSVNAPENGVTENNDFVFGYTQEEAKKNAMHNKQRTSGFSGFGE 64
DB 486 NYRY-----NNNNNN-----NNNNNNNNNNNNNNNNNN 517

QY 65 ENGKG-----HNGALNLFNGKSAQNRFLLTGTNTNLNGKISV 101
DB 518 GNGGNGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 571

RESULT 15
ID KSPL_YEAST STANDARD; PRT; 1029 AA.
AC P38691;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SERINE/THREONINE-PROTEIN KINASE KSPL (EC 2.7.1.-).
GN KSPL OR YHR082C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M335 /2A;
RX MEDLINE=96194625; PubMed=8676864;
RA Fleischmann M., Stagljar I., Aebi M.;
RT "Allele-specific suppression of a Saccharomyces cerevisiae prp20
RT mutation by overexpression of a nuclear serine/threonine protein

kinase."; Mol. Gen. Genet. 250:614-625(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L., Jier M., Johnston L., Langston Y.,
RA Lataille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,
RA Nham M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII."; Science 265:2077-2082(1994).
RL Science 265:2077-2082(1994).
CC -!- FUNCTION: MAY ACT ON PRP20.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-
CC PROTEIN KINASES. BELONGS TO THE CK-II SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X80329; CAA56578.1; -
DR EMBL: U10556; AAB68896.1; -
DR PIR: S46818; S46818.
DR HSSP: Q63450; 1A06.
DR SGD: S0001124; KSPL.
DR InterPro: IPR000719; -
DR InterPro: IPR002290; -
DR Pfam: PF00069; pkinase; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Nuclear protein.
FT DOMAIN 18 351 PROTEIN KINASE.
FT NP_BIND 27 35 ATP (BY SIMILARITY).
FT BINDING 47 47 ATP (BY SIMILARITY).
FT ACT_SITE 207 207 BY SIMILARITY.
SQ SEQUENCE 1029 AA; 117081 MW; 99E2A1FC7CEB3D5A CRC64;

Query Match 13.3%; Score 74.5; DB 1; Length 1029;
Best Local Similarity 27.6%; Pred. No. 6.7;
Matches 24; Conservative 11; Mismatches 35; Indels 17; Gaps 3;

QY 25 GVTENNDFVFGYTQEEAKKNAMHNKQRTSGFSGFGEENGKNGALNLFN----- 79
DB 537 GFSNNNN---KQYRQNRNNNNNNNNNNNNHGSNNYNNF---NNGNSYIKGWNKNKRRP 590

QY 80 -----GKSAQNRFLLTGTNTNLNGKIS 100
DB 591 SSSSYTGKSPLSRYNNMNNNNSSIN 617

Search completed: May 1, 2001, 15:03:35
Job time: 282 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 15:04:51 ; Search time 73.06 seconds
(without alignments)
166.844 Million cell updates/sec

Title: US-09-142-970-2
Perfect score: 562
Sequence: 1 LYXKNRYALKSGSVNAP.....NRLLTGGTNLNGKISVTQG 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL_15:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	562	100.0	993	2 Q9S6X3	Q9S6X3 neisseria m
2	562	100.0	997	2 Q9S6X2	Q9S6X2 neisseria m
3	557	99.1	996	2 Q30574	Q30574 neisseria m
4	546	97.2	996	2 Q57309	Q57309 neisseria m
5	540	96.1	992	2 Q30573	Q30573 neisseria m
6	540	96.1	992	2 Q57035	Q57035 neisseria m
7	540	96.1	992	2 Q9S6X5	Q9S6X5 neisseria m
8	540	96.1	992	2 Q30575	Q30575 neisseria m
9	540	96.1	1561	2 Q51169	Q51169 neisseria m
10	540	96.1	1773	2 Q9JWB9	Q9JWB9 neisseria m
11	536	95.4	1813	2 Q9JWB4	Q9JWB4 neisseria m
12	536	95.4	1813	2 Q9JWB4	Q9JWB4 neisseria m
13	133	23.7	1431	2 Q9JWB4	Q9JWB4 neisseria m
14	115	20.5	1449	2 Q9JWB4	Q9JWB4 neisseria m
15	111.5	19.8	1457	2 Q9X7H1	Q9X7H1 neisseria m
16	111.5	19.8	1457	2 Q9JXL6	Q9JXL6 neisseria m
17	83.5	14.9	319	10 Q41725	Q41725 zinnia eleg
18	83.5	14.9	465	5 Q08212	Q08212 drosophila
19	83	14.8	1577	2 Q55265	Q55265 streptococ

20	82.5	14.7	464	5 Q9VCE3	Q9VCE3 drosophila
21	81	14.4	1225	5 Q9W3A2	Q9W3A2 drosophila
22	80.5	14.3	317	2 Q9R415	Q9R415 neisseria g
23	80.5	14.3	346	2 Q52699	Q52699 neisseria g
24	80.5	14.3	608	5 Q94474	Q94474 dictyostell
25	79.5	14.1	705	2 Q51286	Q51286 neisseria m
26	79.5	14.1	943	10 Q9SIT1	Q9SIT1 arabidopsis
27	79	14.1	586	5 Q9VGH5	Q9VGH5 drosophila
28	77.5	13.8	319	2 Q9RF69	Q9RF69 neisseria g
29	77.5	13.8	319	2 Q9RF57	Q9RF57 neisseria g
30	77.5	13.8	348	2 P95347	P95347 neisseria g
31	77	13.7	163	5 Q9XTW8	Q9XTW8 caenorhabdi
32	77	13.7	712	2 Q9KOV0	Q9KOV0 neisseria m
33	77	13.7	712	2 Q9JPI9	Q9JPI9 neisseria m
34	77	13.7	1650	5 Q77328	Q77328 plasmodium
35	76.5	13.6	631	2 Q48041	Q48041 haemophilus
36	76	13.5	461	10 Q41256	Q41256 nicotiana a
37	76	13.5	604	5 Q02374	Q02374 drosophila
38	75.5	13.4	419	3 Q74786	Q74786 schizosacch
39	75.5	13.4	484	2 Q53394	Q53394 mycobacteri
40	75.5	13.4	702	2 Q69749	Q69749 neisseria m
41	75.5	13.4	764	5 Q96234	Q96234 plasmodium
42	75	13.3	568	5 Q9NL38	Q9NL38 pinctada ma
43	75	13.3	629	2 Q9RSY6	Q9RSY6 deinococcus
44	74.5	13.3	447	5 Q27908	Q27908 pinctada fu
45	74.5	13.3	631	2 Q48043	Q48043 haemophilus

ALIGNMENTS

RESULT 1

Q9S6X3 PRELIMINARY; PRT; 993 AA.

AC Q9S6X3;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE IGAL1 PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G.; Malorny B.; Mueller K.; Seiler A.; Wang J.; del Valle J.; Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30 years of epidemic spread."
RL Mol. Microbiol. 25:1047-1064(1997).
RR EMBL: AF012208; AAC45791.1;
DR INTERPRO: IPR000710;
DR INTERPRO: IPR002195;
DR PRINTS: PR00921; IGASERPTASE.
DR PROSITE: PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 993
SQ SEQUENCE 993 AA; 109441 MW; 109FAA2EF88AC3C6 CRC64;

not p5.2
Sept 1997

Query Match 100.0%; Score 562; DB 2; Length 993;
Best Local Similarity 100.0%; Pred. No. 3e-50;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYXKNRYALKSGSVNAPMPENGVTENNDFVPMGYTQBEAKKNAMHNKNNRISGFSG 60
|||||
Db 558 LYXKNRYALKSGSVNAPMPENGVTENNDFVPMGYTQBEAKKNAMHNKNNRISGFSG 617
|||||

QY 61 FFGKNGKNGHNGALNLFNFKSAQNRLLTGGTNLNGKISVTQG 104
|||||

Db 618 FFGKNGKNGHNGALNLFNFKSAQNRLLTGGTNLNGKISVTQG 661
|||||

```

RESULT 2
Q956X2
ID Q956X2 PRELIMINARY; PRT; 997 AA.
AC Q956X2;
DT 01-MAY-2000 (TREMELrel. 13, Created)
DT 01-MAY-2000 (TREMELrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMELrel. 14, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J., del Valle J.,
RA Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012209; AAC45792.1; -
DR INTERPRO; IPR000710; -
DR INTERPRO; IPR002195; -
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 997
FT NON_TER 997
SQ SEQUENCE 997 AA; 109811 MW; 06F2E361E7E202E0 CRC64;

Query Match 100.0%; Score 562; DB 2; Length 997;
Best Local Similarity 100.0%; Pred. No. 3e-50;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGSVNAPMPENGVTENNNDWFMGYTQEEAKKNAMHNKNQISGFSG 60
DB 558 LYKKNRYRYALKSGSVNAPMPENGVTENNNDWFMGYTQEEAKKNAMHNKNQISGFSG 617
DB 618 FGEENGKGHGALNLFNGKSAQNRFLLTGTGTLNKGKISVTQG 661

RESULT 3
Q30574
ID Q30574 PRELIMINARY; PRT; 996 AA.
AC Q30574;
DT 01-JAN-1998 (TREMELrel. 05, Created)
DT 01-JAN-1998 (TREMELrel. 05, Last sequence update)
DT 01-OCT-2000 (TREMELrel. 15, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J.-F.,
RA del Valle J., Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012206; AAC45789.1; -
DR INTERPRO; IPR000710; -
DR INTERPRO; IPR002195; -
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 996
FT NON_TER 996
SQ SEQUENCE 996 AA; 109688 MW; B3ABDEFE54C337D9 CRC64;

Query Match 97.2%; Score 546; DB 2; Length 996;
Best Local Similarity 97.1%; Pred. No. 1.4e-48;
Matches 101; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGSVNAPMPENGVTENNNDWFMGYTQEEAKKNAMHNKNQISGFSG 60
DB 557 LYKKNRYRYALKSGSVNAPMPENGVTENNNDWFMGYTQEEAKKNAMHNKNQISGFSG 616
DB 617 FGEENGKGHGALNLFNGKSAQNRFLLTGTGTLNKGKISVTQG 660

Query Match 99.1%; Score 557; DB 2; Length 996;
Best Local Similarity 99.0%; Pred. No. 1e-49;
Matches 103; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGSVNAPMPENGVTENNNDWFMGYTQEEAKKNAMHNKNQISGFSG 60
DB 557 LYKKNRYRYALKSGSVNAPMPENGVTENNNDWFMGYTQEEAKKNAMHNKNQISGFSG 616
DB 61 FGEENGKGHGALNLFNGKSAQNRFLLTGTGTLNKGKISVTQG 104
DB 617 FGEENGKGHGALNLFNGKSAQNRFLLTGTGTLNKGKISVTQG 660

RESULT 4
Q57309
ID Q57309 PRELIMINARY; PRT; 996 AA.
AC Q57309;
DT 01-NOV-1996 (TREMELrel. 01, Created)
DT 01-NOV-1998 (TREMELrel. 08, Last sequence update)
DT 01-OCT-2000 (TREMELrel. 15, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J.-F.,
RA del Valle J., Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
RN [2]
RP SEQUENCE OF 37-532 FROM N.A.
RC STRAIN=ETH2;
RX MEDLINE=95302961; PubMed=7783620;
RA Lomholt H., Poulsen K., Mogens K.;
RT "Comparative characterization of the iga gene encoding Iga1 protease
RT in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
RT influenzae.";
RL Mol. Microbiol. 15:495-506(1995).
DR EMBL; AF012207; AAC45790.1; -
DR EMBL; X82469; CAA57852.1; -
DR EMBL; X82468; CAA57851.1; -
DR INTERPRO; IPR000710; -
DR INTERPRO; IPR002195; -
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 996
FT NON_TER 996
SQ SEQUENCE 996 AA; 109717 MW; AB39F7B98A41F986 CRC64;

Query Match 97.2%; Score 546; DB 2; Length 996;
Best Local Similarity 97.1%; Pred. No. 1.4e-48;
Matches 101; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGSVNAPMPENGVTENNNDWFMGYTQEEAKKNAMHNKNQISGFSG 60
DB 557 LYKKNRYRYALKSGSVNAPMPENGVTENNNDWFMGYTQEEAKKNAMHNKNQISGFSG 616
DB 61 FGEENGKGHGALNLFNGKSAQNRFLLTGTGTLNKGKISVTQG 104
DB 617 FGEENGKGHGALNLFNGKSAQNRFLLTGTGTLNKGKISVTQG 660

RESULT 5
Q30573
ID Q30573 PRELIMINARY; PRT; 992 AA.

```

```

AC Q05073;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE IGAI PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98010345; PubMed=9350862;
RA Morelli G., Malorny B., Mueller K., Sella A., Wang J.-F.,
RA del Valle J., Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RL years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012205; AAC45788.1;
DR INTERPRO; IPR00710;
DR INTERPRO; IPR002195;
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 992
SQ SEQUENCE 992 AA; 109255 MW; EF7CA563BC7D7ECB CRC64;

Query Match 96.1%; Score 540; DB 2; Length 992;
Best Local Similarity 96.2%; Pred. No. 5.8e-48;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LYYKNRYALKSGSVNAPENGVTENNNDWVFMGYTQEEAKKNAMHKNQRISGFSG 60
Db 557 LYYKNRYALKSGSVNAPENGVTENNNDWVFMGYTQEEAKKNAMHKNQRISGFSG 616

Qy 61 FFGENGKGHNALNPNFGKSAONRFLLTGNTLNGKISVTQ 104
Db 617 FFGENGKGHNALNPNFGKSAONRFLLTGNTLNGKISVTQ 660

RESULT 6
Q57035 PRELIMINARY; PRT; 992 AA.
AC Q57035;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE IGAI PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98010345; PubMed=9350862;
RA Morelli G., Malorny B., Mueller K., Sella A., Wang J.-F.,
RA del Valle J., Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RL years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012205; AAC45788.1;
DR INTERPRO; IPR00710;
DR INTERPRO; IPR002195;
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 992
SQ SEQUENCE 992 AA; 109255 MW; EF7CA563BC7D7ECB CRC64;

Query Match 96.1%; Score 540; DB 2; Length 992;
Best Local Similarity 96.2%; Pred. No. 5.8e-48;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LYYKNRYALKSGSVNAPENGVTENNNDWVFMGYTQEEAKKNAMHKNQRISGFSG 60
Db 557 LYYKNRYALKSGSVNAPENGVTENNNDWVFMGYTQEEAKKNAMHKNQRISGFSG 616

Qy 61 FFGENGKGHNALNPNFGKSAONRFLLTGNTLNGKISVTQ 104
Db 617 FFGENGKGHNALNPNFGKSAONRFLLTGNTLNGKISVTQ 660

RESULT 7
Q57035 PRELIMINARY; PRT; 992 AA.
AC Q57035;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE IGAI PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RX Morelli G., Malorny B., Mueller K., Sella A., Wang J., del Valle J.,
RA Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RL years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012203; AAC45786.1;
DR INTERPRO; IPR00710;
DR INTERPRO; IPR002195;
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 992
SQ SEQUENCE 992 AA; 109239 MW; 306FF451B7213D49 CRC64;

Query Match 96.1%; Score 540; DB 2; Length 992;
Best Local Similarity 96.2%; Pred. No. 5.8e-48;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LYYKNRYALKSGSVNAPENGVTENNNDWVFMGYTQEEAKKNAMHKNQRISGFSG 60
Db 557 LYYKNRYALKSGSVNAPENGVTENNNDWVFMGYTQEEAKKNAMHKNQRISGFSG 616

Qy 61 FFGENGKGHNALNPNFGKSAONRFLLTGNTLNGKISVTQ 104
Db 617 FFGENGKGHNALNPNFGKSAONRFLLTGNTLNGKISVTQ 660

RESULT 8
Q57035 PRELIMINARY; PRT; 992 AA.
AC Q57035;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)

```

```

DR INTERPRO; IPR00710;
DR INTERPRO; IPR002195;
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 992
SQ SEQUENCE 992 AA; 109228 MW; 3677DD84CE6D9F69 CRC64;

Query Match 96.1%; Score 540; DB 2; Length 992;
Best Local Similarity 96.2%; Pred. No. 5.8e-48;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LYYKNRYALKSGSVNAPENGVTENNNDWVFMGYTQEEAKKNAMHKNQRISGFSG 60
Db 557 LYYKNRYALKSGSVNAPENGVTENNNDWVFMGYTQEEAKKNAMHKNQRISGFSG 616

Qy 61 FFGENGKGHNALNPNFGKSAONRFLLTGNTLNGKISVTQ 104
Db 617 FFGENGKGHNALNPNFGKSAONRFLLTGNTLNGKISVTQ 660

RESULT 7
Q57035 PRELIMINARY; PRT; 992 AA.
AC Q57035;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE IGAI PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RX Morelli G., Malorny B., Mueller K., Sella A., Wang J., del Valle J.,
RA Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RL years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012203; AAC45786.1;
DR INTERPRO; IPR00710;
DR INTERPRO; IPR002195;
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 992
SQ SEQUENCE 992 AA; 109239 MW; 306FF451B7213D49 CRC64;

Query Match 96.1%; Score 540; DB 2; Length 992;
Best Local Similarity 96.2%; Pred. No. 5.8e-48;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LYYKNRYALKSGSVNAPENGVTENNNDWVFMGYTQEEAKKNAMHKNQRISGFSG 60
Db 557 LYYKNRYALKSGSVNAPENGVTENNNDWVFMGYTQEEAKKNAMHKNQRISGFSG 616

Qy 61 FFGENGKGHNALNPNFGKSAONRFLLTGNTLNGKISVTQ 104
Db 617 FFGENGKGHNALNPNFGKSAONRFLLTGNTLNGKISVTQ 660

RESULT 8
Q57035 PRELIMINARY; PRT; 992 AA.
AC Q57035;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)

```

```
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Sella A., Wang J., del Valle J.,
  Ahtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
  years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL: AF012204; AAC45787.1; -.
DR INTERPRO: IPR000710; -.
DR INTERPRO: IPR002195; -.
DR PRINTS: PR00921; IGASERPTASE.
DR PROSITE: PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 992
FT SEQUENCE 992 AA; 109239 MW; 9ED0A41019917CE3 CRC64;

Query Match 96.1%; Score 540; DB 2; Length 992;
Best Local Similarity 96.2%; Pred. No. 5.8e-48;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGSVNAPMPENGVTENNNDWVFMGYTQEEAKKNAMHNKRNQISGFSG 60
DB 557 LYKKNRYRYALKSGSVNAPMPENGVTENNNDWVFMGYTQEEAKKNAMHNKRNQISGFSG 616

QY 61 FFGENGKGHGALNLFNGKSAQNRFLLTGTGTLNKGKISVTQ 104
DB 617 FFGENGKGHGALNLFNGKSAQNRFLLTGTGTLNKGKISVTQ 660

RESULT 9
O30575 PRELIMINARY; PRT; 997 AA.
AC O30575;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Sella A., Wang J.-F.,
  del Valle J., Ahtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
  years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL: AF012210; AAC45793.1; -.
DR INTERPRO: IPR000710; -.
DR INTERPRO: IPR002195; -.
DR PRINTS: PR00921; IGASERPTASE.
DR PROSITE: PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 997
FT SEQUENCE 997 AA; 110152 MW; 4DE40DE594E5E28E CRC64;

Query Match 96.1%; Score 540; DB 2; Length 997;
Best Local Similarity 96.2%; Pred. No. 5.9e-48;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGSVNAPMPENGVTENNNDWVFMGYTQEEAKKNAMHNKRNQISGFSG 60
DB 557 LYKKNRYRYALKSGSVNAPMPENGVTENNNDWVFMGYTQEEAKKNAMHNKRNQISGFSG 616
```

```
DB 558 LYKKNRYRYALKSGSVNAPMPENGVTENNNDWVFMGYTQEEAKKNAMHNKRNQISGFSG 617
QY 61 FFGENGKGHGALNLFNGKSAQNRFLLTGTGTLNKGKISVTQ 104
DB 618 FFGENGKGHGALNLFNGKSAQNRFLLTGTGTLNKGKISVTQ 661

RESULT 10
O51169 PRELIMINARY; PRT; 1561 AA.
AC O51169;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE IGAL PROTEASE.
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Lomholt H., Poulsen K., Mogens K.;
RT "Comparative characterization of the iga gene encoding Iga1 protease
  in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
  influenzae.";
RL Mol. Microbiol. 15:495-506(1995).
DR EMBL: X82474; CAA57857.1; -.
DR MEROPS: S06.001; -.
DR INTERPRO: IPR000710; -.
DR INTERPRO: IPR002195; -.
DR PRINTS: PR00921; IGASERPTASE.
DR PROSITE: PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
SQ SEQUENCE 1561 AA; 171848 MW; 1C96E291A00017D5 CRC64;

Query Match 96.1%; Score 540; DB 2; Length 1561;
Best Local Similarity 96.2%; Pred. No. 9.9e-48;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGSVNAPMPENGVTENNNDWVFMGYTQEEAKKNAMHNKRNQISGFSG 60
DB 584 LYKKNRYRYALKSGSVNAPMPENGVTENNNDWVFMGYTQEEAKKNAMHNKRNQISGFSG 643

QY 61 FFGENGKGHGALNLFNGKSAQNRFLLTGTGTLNKGKISVTQ 104
DB 644 FFGENGKGHGALNLFNGKSAQNRFLLTGTGTLNKGKISVTQ 687

RESULT 11
O9JVB9 PRELIMINARY; PRT; 1773 AA.
AC O9JVB9;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE IGAL PROTEASE (EC 3.4.21.7).
GN IGA OR NMA0905.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=22451 / SEROGROUP A / SEROTYPE 4A;
RT MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
  Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
  Davies S.R., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
  Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
  Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
  Whitehead S., Spratt B.G., Barrell B.G.;
```

RT *Complete DNA sequence of a serogroup A strain of Neisseria
RL Nature 404:502-506(2000).
DR EMBL: AL162754; CAB84182.1; -
KW Protease; Hydrolase.
SQ SEQUENCE 1773 AA; 196350 MW; CAC19E7313D76CE1 CRC64;

Query Match 96.1%; Score 540; DB 2; Length 1773;
Best Local Similarity 96.2%; Pred. No. 1.2e-47;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Oy 1 LYXNYRYALKSGSVNAPNGVTENNDDWFMGYTQEEAKKNAMHKNQRISGFSG 60
Db 573 LYXNYRYALKSGSVNAPNGVTENNDDWFMGYTQEEAKKNAMHKNQRISGFSG 632

Oy 61 FFGENGKGHGNALNPNFGSAQRNELLTGGTNLNGKISVTQ 104
Db 633 FFGENGKGHGNALNPNFGSAQRNELLTGGTNLNGKISVTQ 676

RESULT 12
Oy 633 FFGENGKGHGNALNPNFGSAQRNELLTGGTNLNGKISVTQ 676
ID Q9K0B4 PRELIMINARY; PRT; 1815 AA.
AC Q9K0B4;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE IGA-SPECIFIC SERINE ENDOPEPTIDASE.
GN NMB0700.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citti H., Vamathevan J.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT *Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58.*;
RL Science 287:1809-1815(2000).
DR EMBL: AE002424; AAF41117.1; -
DR TIGR: NMB0700; -
SQ SEQUENCE 1815 AA; 201077 MW; 2259D4D71762C57F CRC64;

Query Match 95.4%; Score 536; DB 2; Length 1815;
Best Local Similarity 95.2%; Pred. No. 3.1e-47;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 1 LYXNYRYALKSGSVNAPNGVTENNDDWFMGYTQEEAKKNAMHKNQRISGFSG 60
Db 573 LYXNYRYALKSGSVNAPNGVTENNDDWFMGYTQEEAKKNAMHKNQRISGFSG 632
Oy 61 FFGENGKGHGNALNPNFGSAQRNELLTGGTNLNGKISVTQ 104
Db 633 FFGENGKGHGNALNPNFGSAQRNELLTGGTNLNGKISVTQ 676

RESULT 13
Oy 633 FFGENGKGHGNALNPNFGSAQRNELLTGGTNLNGKISVTQ 676
ID Q9JXK3 PRELIMINARY; PRT; 1431 AA.
AC Q9JXK3;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)

DE SERINE-TYPE PEPTIDASE.
GN NMB1998.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citti H., Vamathevan J.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Masignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT *Complete genome sequence of Neisseria meningitidis serogroup B strain
MC58.*;
RL Science 287:1809-1815(2000).
DR EMBL: AE002549; AAF42325.1; -
DR TIGR: NMB1998; -
SQ SEQUENCE 1431 AA; 157645 MW; 6801F0E209C24EFC CRC64;

Query Match 23.7%; Score 133; DB 2; Length 1431;
Best Local Similarity 26.3%; Pred. No. 2.1e-05;
Matches 36; Conservative 22; Mismatches 43; Indels 36; Gaps 5;
Oy 3 YKNYR--YYALKSGSVNAPNGVTENNDDWFMGYTQEEAKKNAMHKN--NORISGFS 59
Db 555 HNRRTDYFILKPGGNPREFFPLN-MKNSTSWQFIGNRRQAAEQVAENAPDLITFG 613
Oy 60 GFFGE--ENGK-----HNGALNLPNGSAQRNF 87
Db 614 GYLGENAQTGKAAPSYKTNAAIEKTRHIANAAVYGRPEYRYNGALNHYRKRDTSTL 673
Oy 88 LLTGTLNGLKISVTQ 104
Db 674 LLNGMNLNGEVLEGG 690

RESULT 14
Oy 88 LLTGTLNGLKISVTQ 104
ID Q9JWB4 PRELIMINARY; PRT; 1449 AA.
AC Q9JWB4;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TremBLrel. 15, Last annotation update)
DE IGA-SPECIFIC SERINE ENDOPEPTIDASE (EC 3.4.21.72).
GN IGA2 OR NMA0457.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT *Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491.*;
RL Nature 404:502-506(2000).
DR EMBL: AL162753; CAB83754.1; -
KW Hydrolase.
SQ SEQUENCE 1449 AA; 159070 MW; B858D8854BE50146 CRC64;

Query Match 20.5%; Score 115; DB 2; Length 1449;

992

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 15:04:52 ; Search time 73.06 seconds
(without alignments)
166.844 Million cell updates/sec

Title: US-09-142-970-3
Perfect score: 563
Sequence: 1 LYKNRYALKSGSVNAP.....NRELLTGTLNKGKISVTQG 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_15:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_unclassified:*
- 13: sp_vertebrate:*
- 14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	563	100.0	996	2	Q57309 neisseria m
2	546	97.0	993	2	Q9S6X3 neisseria m
3	546	97.0	997	2	Q9S6X2 neisseria m
4	541	96.1	996	2	Q30574 neisseria m
5	537	95.4	992	2	Q30573 neisseria m
6	537	95.4	992	2	Q57035 neisseria m
7	537	95.4	992	2	Q9S6X5 neisseria m
8	537	95.4	992	2	Q9S6X4 neisseria m
9	537	95.4	997	2	Q30575 neisseria m
10	537	95.4	1561	2	Q51169 neisseria m
11	537	95.4	1773	2	Q9JWB9 neisseria m
12	533	94.7	1815	2	Q9K0B4 neisseria m
13	132	23.4	1431	2	Q9JXK3 neisseria m
14	112.5	20.0	1449	2	Q9JWB4 neisseria m
15	109	19.4	1457	2	Q9X7H1 neisseria m
16	109	19.4	1457	2	Q9JXL6 neisseria m
17	84	14.9	1225	5	Q9WJA2 drosophila
18	82.5	14.7	319	10	Q41725 zinnia eleg
19	80	14.2	163	5	Q9XTW8 caenorhabdi

20	80	14.2	456	5	Q20936
21	80	14.2	2529	2	Q25579
22	79	14.0	179	5	O02233
23	79	14.0	402	5	O46309
24	79	14.0	629	2	Q9RSY6
25	79	14.0	1650	5	O77328
26	79	14.0	2399	2	Q9ZKS9
27	78.5	13.9	447	5	Q27908
28	78	13.9	461	10	Q41256
29	78	13.9	568	5	Q9NL38
30	78	13.9	1577	2	Q55265
31	77.5	13.8	718	10	Q9ZUW8
32	77.5	13.8	943	10	Q9SIT1
33	77	13.7	162	2	Q9RG12
34	77	13.7	214	2	Q9XDI9
35	77	13.7	3032	5	O97278
36	76.5	13.6	419	3	O74786
37	76	13.5	586	5	Q9VGH5
38	76	13.5	604	5	O02374
39	75.5	13.4	317	2	Q9R415
40	75.5	13.4	346	2	O52699
41	75.5	13.4	465	5	O08212
42	75.5	13.4	484	2	O53394
43	75	13.3	697	5	Q20220
44	74.5	13.2	464	5	Q9VCE3
45	74.5	13.2	1366	5	Q9V6W8

ALIGNMENTS

RESULT 1

Q57309 PRELIMINARY; PRT; 996 AA.

AC Q57309; 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)

DE IGAL PROTEASE PRECURSOR (FRAGMENT).

GN IGA.

OS Neisseria meningitidis.

OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.

NCBI_TaxID=487;

SEQUENCE FROM N.A.

MEDLINE=98010345; PubMed=9350862;

Morelli G., Malorny B., Mueller K., Seiler A., Wang J.-F.,

del Valle J., Achtman M.;

"Clonal descent and microevolution of Neisseria meningitidis during 30

years of epidemic spread."

Mol. Microbiol. 25:1047-1064(1997).

SEQUENCE OF 37-532 FROM N.A.

STRAIN=ETH2;

MEDLINE=95302961; PubMed=7783620;

Lomholt H., Poulsen K., Mogens K.;

"Comparative characterization of the iga gene encoding Igal protease

in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus

influenzae."

Mol. Microbiol. 15:495-506(1995).

EMBL; AF012207; AAC45790.1; -

EMBL; X82469; CAA57852.1; -

EMBL; X82468; CAA57851.1; -

INTERPRO; IPR000710; -

INTERPRO; IPR002195; -

PRINTS; PR00921; IGASERPTASE.

PROSITE; PS00482; DIHYDROOCTASE_1; UNKNOWN_1.

Protease.

NON_TER 1

FT 996

SEQUENCE 996 AA; 109717 MW; AB39F7B98A41F986 CRC64;

```
Query Match 100.0%; Score 563; DB 2; Length 996;
Best Local Similarity 100.0%; Pred. No. 7e-51;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGGSVNAPMPENGQOTENNDFVFMGYKQEEAKKNAMHNKNQRISGFSG 60
DB 557 LYKKNRYRYALKSGGSVNAPMPENGQOTENNDFVFMGYKQEEAKKNAMHNKNQRISGFSG 616
QY 61 FFEENGKGHGALNLFNGKSAQNRFLLTGTTNLNGKISVTQ 104
DB 617 FFEENGKGHGALNLFNGKSAQNRFLLTGTTNLNGKISVTQ 660

RESULT 2
QY 956X3 PRELIMINARY; PRT; 993 AA.
AC Q956X3:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Sella A., Wang J., del Valle J.,
RA Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012208; AAC45792.1; -.
DR INTERPRO; IPR000710; -.
DR INTERPRO; IPR002195; -.
DR PRINTS; PRO0921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 993
FT SEQUENCE 993 AA; 109441 MW; 109FAA2EF88AC3C6 CRC64;

Query Match 97.0%; Score 546; DB 2; Length 993;
Best Local Similarity 97.1%; Pred. No. 4.2e-49;
Matches 101; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGGSVNAPMPENGQOTENNDFVFMGYKQEEAKKNAMHNKNQRISGFSG 60
DB 558 LYKKNRYRYALKSGGSVNAPMPENGQOTENNDFVFMGYTQEEAKKNAMHNKNQRISGFSG 617
QY 61 FFEENGKGHGALNLFNGKSAQNRFLLTGTTNLNGKISVTQ 104
DB 618 FFEENGKGHGALNLFNGKSAQNRFLLTGTTNLNGKISVTQ 661

RESULT 3
QY 956X2 PRELIMINARY; PRT; 997 AA.
AC Q956X2:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Sella A., Wang J., del Valle J.,
RA Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
```

```
RT Years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012209; AAC45792.1; -.
DR INTERPRO; IPR000710; -.
DR INTERPRO; IPR002195; -.
DR PRINTS; PRO0921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 997
FT SEQUENCE 997 AA; 109811 MW; 06F2E361E7E202E0 CRC64;

Query Match 97.0%; Score 546; DB 2; Length 997;
Best Local Similarity 97.1%; Pred. No. 4.3e-49;
Matches 101; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGGSVNAPMPENGQOTENNDFVFMGYKQEEAKKNAMHNKNQRISGFSG 60
DB 558 LYKKNRYRYALKSGGSVNAPMPENGQOTENNDFVFMGYTQEEAKKNAMHNKNQRISGFSG 617
QY 61 FFEENGKGHGALNLFNGKSAQNRFLLTGTTNLNGKISVTQ 104
DB 618 FFEENGKGHGALNLFNGKSAQNRFLLTGTTNLNGKISVTQ 661

RESULT 4
QY 9574 PRELIMINARY; PRT; 996 AA.
AC Q9574:
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98010345; PubMed=9350862;
RA Morelli G., Malorny B., Mueller K., Sella A., Wang J., F.,
RA del Valle J., Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012206; AAC45789.1; -.
DR INTERPRO; IPR000710; -.
DR INTERPRO; IPR002195; -.
DR PRINTS; PRO0921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 996
FT SEQUENCE 996 AA; 109688 MW; B3ABDEFE54C337D9 CRC64;

Query Match 96.1%; Score 541; DB 2; Length 996;
Best Local Similarity 96.2%; Pred. No. 1.4e-48;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGGSVNAPMPENGQOTENNDFVFMGYKQEEAKKNAMHNKNQRISGFSG 60
DB 557 LYKKNRYRYALKSGGSVNAPMPENGQOTENNDFVFMGYTQEEAKKNAMHNKNQRISGFSG 616
QY 61 FFEENGKGHGALNLFNGKSAQNRFLLTGTTNLNGKISVTQ 104
DB 617 FFEENGKGHGALNLFNGKSAQNRFLLTGTTNLNGKISVTQ 660

RESULT 5
QY 9573 PRELIMINARY; PRT; 992 AA.
AC Q9573:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2000 (Tremblrel. 14, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Sella A., Wang J., del Valle J.,
RA Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
```



```

AC O=0573;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98010345; PubMed=9350862;
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J.-F.,
   del Valle J., Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
   years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012205; AAC45786.1; -
DR INTERPRO: IPR000710; -
DR PRINTS; PR002195; -
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT SEQUENCE 992 AA; 109255 MW; EF7CA563BC7D7ECB CRC64;
SQ

Query Match 95.4%; Score 537; DB 2; Length 992;
Best Local Similarity 95.2%; Pred. No. 3.7e-48;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGGSVNAPMPENGOTENNDDWFMGYKQEEAKKNAMHNKNORISGFSG 60
Db 557 LYKKNRYRYALKSGGSVNAPMPENGOTENNDDWFMGYKQEEAKKNAMHNKNORISGFSG 616

QY 61 FFGENGKGHGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 104
Db 617 FFGENGKGHGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 660

RESULT 6
Q57035 ID Q57035 PRELIMINARY; PRT; 992 AA.
AC Q57035;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98010345; PubMed=9350862;
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J.-F.,
   del Valle J., Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
   years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012211; AAC45794.1; -
DR INTERPRO: IPR000710; -
DR PRINTS; PR002195; -
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT SEQUENCE 992 AA; 109255 MW; EF7CA563BC7D7ECB CRC64;
SQ

Query Match 95.4%; Score 537; DB 2; Length 992;
Best Local Similarity 95.2%; Pred. No. 3.7e-48;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGGSVNAPMPENGOTENNDDWFMGYKQEEAKKNAMHNKNORISGFSG 60
Db 557 LYKKNRYRYALKSGGSVNAPMPENGOTENNDDWFMGYKQEEAKKNAMHNKNORISGFSG 616

QY 61 FFGENGKGHGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 104
Db 617 FFGENGKGHGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 660

RESULT 7
Q57035 ID Q57035 PRELIMINARY; PRT; 992 AA.
AC Q57035;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J., del Valle J.,
   Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
   years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012203; AAC45786.1; -
DR INTERPRO: IPR000710; -
DR PRINTS; PR002195; -
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT SEQUENCE 992 AA; 109239 MW; 306FF451B7213D49 CRC64;
SQ

Query Match 95.4%; Score 537; DB 2; Length 992;
Best Local Similarity 95.2%; Pred. No. 3.7e-48;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGGSVNAPMPENGOTENNDDWFMGYKQEEAKKNAMHNKNORISGFSG 60
Db 557 LYKKNRYRYALKSGGSVNAPMPENGOTENNDDWFMGYKQEEAKKNAMHNKNORISGFSG 616

QY 61 FFGENGKGHGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 104
Db 617 FFGENGKGHGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 660

RESULT 8
Q57035 ID Q57035 PRELIMINARY; PRT; 992 AA.
AC Q57035;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

```

```

DR INTERPRO: IPR000710; -
DR INTERPRO: IPR002195; -
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT SEQUENCE 992 AA; 109228 MW; 3677DDE4CE6D9F69 CRC64;
SQ

Query Match 95.4%; Score 537; DB 2; Length 992;
Best Local Similarity 95.2%; Pred. No. 3.7e-48;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGGSVNAPMPENGOTENNDDWFMGYKQEEAKKNAMHNKNORISGFSG 60
Db 557 LYKKNRYRYALKSGGSVNAPMPENGOTENNDDWFMGYKQEEAKKNAMHNKNORISGFSG 616

QY 61 FFGENGKGHGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 104
Db 617 FFGENGKGHGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 660

RESULT 7
Q57035 ID Q57035 PRELIMINARY; PRT; 992 AA.
AC Q57035;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE IGAL PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Seiler A., Wang J., del Valle J.,
   Achtman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
   years of epidemic spread.";
RL Mol. Microbiol. 25:1047-1064(1997).
DR EMBL; AF012203; AAC45786.1; -
DR INTERPRO: IPR000710; -
DR PRINTS; PR002195; -
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT SEQUENCE 992 AA; 109239 MW; 306FF451B7213D49 CRC64;
SQ

Query Match 95.4%; Score 537; DB 2; Length 992;
Best Local Similarity 95.2%; Pred. No. 3.7e-48;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKKNRYRYALKSGGSVNAPMPENGOTENNDDWFMGYKQEEAKKNAMHNKNORISGFSG 60
Db 557 LYKKNRYRYALKSGGSVNAPMPENGOTENNDDWFMGYKQEEAKKNAMHNKNORISGFSG 616

QY 61 FFGENGKGHGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 104
Db 617 FFGENGKGHGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 660

RESULT 8
Q57035 ID Q57035 PRELIMINARY; PRT; 992 AA.
AC Q57035;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)

```

DE IGA1 PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Sella A., Wang J., del Valle J.,
RA Achman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread";
RL Mol. Microbiol. 25:1047-1064 (1997).
DR EMBL; AF012204; AAC45787.1; -;
DR INTERPRO; IPR00710; -;
DR INTERPRO; IPR002195; -;
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 992 992
SQ SEQUENCE 992 AA; 109239 MW; 9ED0A41019917CE3 CRC64;

Query Match 95.4%; Score 537; DB 2; Length 992;
Best Local Similarity 95.2%; Pred. No. 3.7e-48;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKKNRYALKSGSVNAPMPENGQTNNDWVFMGYKQEEAKKNAMHNKNNRISGFSG 60
|||||
DB 557 LYKKNRYALKSGSVNAPMPENGQTNNDWVFMGYKQEEAKKNAMHNKNNRISGFSG 616
|||||

QY 61 FFEENGKGHNALNFGKSAQNRFLLTGCTNLNGKISVTQ 104
|||||
DB 617 FFEENGKGHNALNFGKSAQNRFLLTGCTNLNGKISVTQ 660
|||||

RESULT 9
O30575 PRELIMINARY; PRT; 997 AA.
AC O30575;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE IGA1 PROTEASE PRECURSOR (FRAGMENT).
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RA Morelli G., Malorny B., Mueller K., Sella A., Wang J.-F.,
RA del Valle J., Achman M.;
RT "Clonal descent and microevolution of Neisseria meningitidis during 30
RT years of epidemic spread";
RL Mol. Microbiol. 25:1047-1064 (1997).
DR EMBL; AF012210; AAC45793.1; -;
DR INTERPRO; IPR00710; -;
DR INTERPRO; IPR002195; -;
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
FT NON_TER 1
FT NON_TER 997 997
SQ SEQUENCE 997 AA; 110152 MW; 4DE40DE594E5E28E CRC64;

Query Match 95.4%; Score 537; DB 2; Length 997;
Best Local Similarity 95.2%; Pred. No. 3.8e-48;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKKNRYALKSGSVNAPMPENGQTNNDWVFMGYKQEEAKKNAMHNKNNRISGFSG 60
|||||

DB 558 LYKKNRYALKSGSVNAPMPENGQTNNDWVFMGYKQEEAKKNAMHNKNNRISGFSG 617
QY 61 FFEENGKGHNALNFGKSAQNRFLLTGCTNLNGKISVTQ 104
|||||
DB 618 FFEENGKGHNALNFGKSAQNRFLLTGCTNLNGKISVTQ 661
|||||

RESULT 10
Q51169 PRELIMINARY; PRT; 1561 AA.
AC Q51169;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE IGA1 PROTEASE.
GN IGA.
OS Neisseria meningitidis.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=487;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HF13;
RX MEDLINE=95302961; PubMed=7783620;
RA Lomholt H., Foulson K., Mogens K.;
RT "Comparative characterization of the iga gene encoding Iga1 protease
RT in Neisseria meningitidis, Neisseria gonorrhoeae and Haemophilus
RT influenzae";
RL Mol. Microbiol. 15:495-506 (1995).
DR EMBL; X82474; CAA57857.1; -;
DR MEROPS; S06.001; -;
DR INTERPRO; IPR000710; -;
DR INTERPRO; IPR002195; -;
DR PRINTS; PR00921; IGASERPTASE.
DR PROSITE; PS00482; DIHYDROOROTASE_1; UNKNOWN_1.
KW Protease.
SQ SEQUENCE 1561 AA; 171848 MW; 1C96E291A00017D5 CRC64;

Query Match 95.4%; Score 537; DB 2; Length 1561;
Best Local Similarity 95.2%; Pred. No. 6.3e-48;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LYKKNRYALKSGSVNAPMPENGQTNNDWVFMGYKQEEAKKNAMHNKNNRISGFSG 60
|||||
DB 584 LYKKNRYALKSGSVNAPMPENGQTNNDWVFMGYKQEEAKKNAMHNKNNRISGFSG 643
|||||

QY 61 FFEENGKGHNALNFGKSAQNRFLLTGCTNLNGKISVTQ 104
|||||
DB 644 FFEENGKGHNALNFGKSAQNRFLLTGCTNLNGKISVTQ 687
|||||

RESULT 11
Q9JVB9 PRELIMINARY; PRT; 1773 AA.
AC Q9JVB9;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)
DE IGA1 PROTEASE (EC 3.4.21.7).
GN IGA OR NMA0905.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;

AC	Q9JXR3	DT	DT	DT	DT	DT
		01-OCT-2000	(TReMBLrel. 15, Created)			
		01-OCT-2000	(TReMBLrel. 15, Last sequence update)			
		01-OCT-2000	(TReMBLrel. 15, Last annotation update)			

Query Match 20.0%; Score 112.5; DB 2; Length 1449;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 15:03:37 ; Search time 26.69 seconds
(without alignments)
133.480 Million cell updates/sec

Title: US-09-142-970-4

Perfect score: 569

Sequence: 1 LYKRYRYALKSGRLNAP.....NRFLTGGANLNGNGRVPK 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	523	91.9	1532	1 IGA_NEIGO	P09790 neisseria g
2	281.5	49.5	1541	1 IGA_HAEIN	P42782 haemophilus
3	281.5	49.5	1694	1 IGAO_HAEIN	P44969 haemophilus
4	281.5	49.5	1702	1 IGA2_HAEIN	P45384 haemophilus
5	270.5	47.5	1545	1 IGA3_HAEIN	P45385 haemophilus
6	270.5	47.5	1849	1 IGA4_HAEIN	P45386 haemophilus
7	91	16.0	1394	1 HAP_HAEIN	P45387 haemophilus
8	89	15.6	1409	1 HAP1_HAEIN	P44596 haemophilus
9	79.5	14.0	451	1 ARP2_PLAFA	P13824 plasmodium
10	73.5	12.9	1596	1 MAM_DROME	P21519 drosophila
11	72	12.7	1029	1 KSP1_YEAST	P38691 saccharomyc
12	71.5	12.6	1254	1 POLS_EEVEV8	P05674 venezuelan
13	71.5	12.6	1254	1 POLS_EEVEE	P36330 venezuelan
14	71.5	12.6	1254	1 POLS_EEVEV1	P09592 venezuelan
15	71.5	12.6	1255	1 POLS_EEVP3	P36332 venezuelan
16	71.5	12.6	1255	1 POLS_EEVP	P36333 venezuelan
17	71.5	12.6	1538	1 GLSF_PORPU	P51375 porphyra pu
18	71	12.5	537	1 ARP_PLAFA	P04931 plasmodium
19	71	12.5	700	1 NONA_DROME	Q04047 drosophila
20	70.5	12.4	719	1 ARP_YEAST	P32770 saccharomyc
21	70	12.3	534	1 NABA_YEAST	Q99383 saccharomyc
22	70	12.3	592	1 RB56_HUMAN	Q92804 homo sapien
23	70	12.3	808	1 GCS1_SCHPO	O14255 schizosacch
24	70	12.3	989	1 PTP3_DICDI	P54637 dictyosteli
25	69.5	12.2	490	1 CAR3_DICDI	P35352 dictyosteli
26	69.5	12.2	954	1 XYN4_RUMFL	P29126 ruminoococu
27	68.5	12.0	522	1 IKAR_ONCMY	O13089 oncorhynchu
28	68	12.0	452	1 PUB1_YEAST	P32588 saccharomyc
29	68	12.0	1585	1 P3K3_DICDI	P54675 dictyosteli
30	67.5	11.9	341	1 OMPU_VIBCH	P97085 vibrio chol
31	67.5	11.9	599	1 BAL_MOUSE	Q64285 mus musculu
32	67.5	11.9	801	1 DHGA_ACICA	P05465 acinetobact
33	67	11.8	430	1 SHU7_ECOLI	P09746 escherichia

RESULT 1

ID	IGA_NEIGO	STANDARD;	PRT;	1532 AA.
AC	P09790;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-MAR-1989 (Rel. 10, Last sequence update)			
DE	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	IGA-SPECIFIC SERINE ENDOPEPTIDASE PRECURSOR (EC 3.4.21.72) (IGA			
DE	PROTEASE).			
GN	IGA.			
OS	Neisseria gonorrhoeae.			
OC	Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.			
OX	NCBI_TaxID=485;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RC	STRAIN=MS11.			
RX	MEDLINE=87115823; PubMed=3027577;			
RA	Pohlner J., Halter R., Beyreuther K., Meyer T.F.;			
RT	"Gene structure and extracellular secretion of Neisseria gonorrhoeae			
RT	Iga protease.";			
RL	Nature 325:458-462(1987).			
RN	[2]			
RP	ACTIVE SITE			
RX	MEDLINE=90154052; PubMed=2105953;			
RA	Bachovchin W.W., Plaut A.G., Flentke G.R., Lynch M., Kettner C.A.;			
RT	"Inhibition of IgA1 proteinases from Neisseria gonorrhoeae and			
RT	Hemophilus influenzae by peptide prolyl boronic acids.";			
J. Biol. Chem.	265:3738-3743(1990).			
CC	-I- FUNCTION: THIS PROTEASE IS SPECIFIC FOR IMMUNOGLOBULIN A.			
CC	-I- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT			
CC	CERTAIN PRO-I-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE			
CC	SUBSTRATES ARE KNOWN.			
CC	-I- SUBCELLULAR LOCATION: SECRETED.			
CC	-I- DOMAIN: THE SIGNAL PEPTIDE GUIDE THE PRECURSOR TO THE PERIPLASMIC			
CC	SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE			
CC	OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE			
CC	DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS.			
CC	-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; X04835; CAA28538.1; -			
DR	PIR; A26039; A26039.			
DR	MEROPS; S06.001; -			
DR	InterPro; IPR000710; -			
DR	PRINTS; PR00921; IGASERPTASE.			
KW	Hydrolase; Serine protease; Zymogen; Autocatalytic cleavage;			
FT	Transmembrane; Signal.			
FT	SIGNAL 1 27			
FT	CHAIN 28 986			
FT	IGA-SPECIFIC SERINE ENDOPEPTIDASE.			

ALIGNMENTS

FT PROPEP 987 1532 1532 1532 HELPER PEPTIDE.
FT ACT_SITE 278 278 278 278 POTENTIAL.
FT SITE 986 987 987 987 CLEAVAGE (AUTO-).
FT SITE 1018 1019 1019 1019 CLEAVAGE (AUTO-).
FT SITE 1121 1122 1122 1122 CLEAVAGE (AUTO-).
SQ SEQUENCE 1532 AA; 168976 MW; 68FF4112BD22F40D CRC64;

Query Match 91.9%; Score 523; DB 1; Length 1532;
Best Local Similarity 88.4%; Pred. No. 3.7e-45;
Matches 99; Conservative 2; Mismatches 1; Indels 10; Gaps 2;

QY 1 LYKKNRYALKSGRLNAPMPENGVAENNDVFMGYTQEEARKNMMNNRRIGDFG 60
DB 584 LYKKNRYALKSGRLNAPMPENGVAENNDVFMGYTQEEARKNMMNNRRIGDFG 643
QY 61 FFEDENGKHGALNLFNGKSAQNRFLTTGGANLNG-----GN-----GRP 102
DB 644 FFEDENGKHGALNLFNGKSAQNRFLTTGGANLNGKISVTQGNVLLSGRP 695

RESULT 2
IGAL_HAEIN STANDARD; PRT; 1541 AA.
ID IGAL_HAEIN STANDARD; PRT; 1541 AA.
AC P42782;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (EC 3.4.21.72) (IGAL PROTEASE).
GN IGA.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HK368 / SEROTYPE B;
RX MEDLINE=89379374; PubMed=2506130;
RA Poulsen K., Brandt J., Hjorth J.P., Thøgersen H.C., Kilian M.;
RT "Cloning and sequencing of the immunoglobulin A1 protease gene (iga)
of Haemophilus influenzae serotype b";
RL Infect. Immun. 57:3097-3105(1989).
RN [2]
RP MUTAGENESIS OF SER-288.
RC STRAIN=HK368 / SEROTYPE B;
RX MEDLINE=9223499; PubMed=1373717;
RA Poulsen K., Reinholdt J., Kilian M.;
RT "A comparative genetic study of serologically distinct Haemophilus
influenzae type 1 immunoglobulin A1 proteases.";
RL J. Bacteriol. 174:2913-2921(1992).
CC -!- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
PRODUCING INTACT FC AND FAB FRAGMENTS.
CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT
CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE
SUBSTRATES ARE KNOWN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; X64357; CAA45708.1; -;
CC EMBL; M87492; AAA24969.1; -;

DR MEROPS; S06.001; -;
DR InterPro; IPR000710; -;
DR PRINTS; PRO0921; IGASERPTASE.
KW Hydrolase; Serine protease; Transmembrane; Zymogen; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 1008 IMMUNOGLOBULIN A1 PROTEASE.
FT PROPEP 1009 1541 HELPER PEPTIDE (POTENTIAL).
FT ACT_SITE 288 288 PROBABLE.
FT MUTAGEN 288 288 S->T: LOSS OF ACTIVITY.
SQ SEQUENCE 1541 AA; 169370 MW; CE7257CB3196C600 CRC64;

Query Match 49.5%; Score 281.5; DB 1; Length 1541;
Best Local Similarity 48.2%; Pred. No. 1e-20;
Matches 54; Conservative 17; Mismatches 30; Indels 11; Gaps 2;

QY 1 LYKKNRYALKSGRLNAPMPENGVAENNDVFMGYTQEEARKNMMNNRRIGDFG 60
DB 595 LLENVYYALKRGASTRSELSPKSGESNENLYMGKTSDEAKRNMNNHNNRNGFNG 654
QY 61 FFEDENGKHGALNLFNGKSAQNRFLTTGGANLNG-----NGRP 102
DB 655 YFGEDEGK-NNGNLNVTFKGSQNRFLTTGGTNLNGDLTVEKFTLFLSGRP 705

RESULT 3
IGAO_HAEIN STANDARD; PRT; 1694 AA.
ID IGAO_HAEIN STANDARD; PRT; 1694 AA.
AC P44969;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (EC 3.4.21.72) (IGAL PROTEASE).
GN IGA OR IGA1 OR HI0990.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SEROTYPE D;
RA Wright A., Fishman Y., Tai F., Plaut A.G.;
RL Submitted (MAY-1991) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd.";
RL Science 269:496-512(1995).
CC -!- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
PRODUCING INTACT FC AND FAB FRAGMENTS.
CC -!- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT
CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE
SUBSTRATES ARE KNOWN.
CC -!- SUBCELLULAR LOCATION: SECRETED.
CC -!- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
SIMILARITY).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X59800; ; NOT_ANNOTATED_CDS.

DR EMBL; U32779; AAC22651.1; ;

DR MEROPS; S06.001; ;

DR TIGR; HI0990; ;

DR InterPro; IPR000710; ;

DR PRINTS; PR00921; IGASERPTASE.

FW HydroLase; Serine protease; Transmembrane; Zymogen; Signal.

FT SIGNAL 1 25 POTENTIAL.

FT CHAIN 26 1014 IMMUNOGLOBULIN A1 PROTEASE.

FT PROPEP 1015 1694 HELPER PEPTIDE (POTENTIAL).

FT ACT_SITE 288 288 PROBABLE.

FT CONFLICT 253 254 EN -> GV (IN REF. 1).

FT CONFLICT 272 272 G -> A (IN REF. 1).

FT CONFLICT 464 464 G -> E (IN REF. 1).

FT CONFLICT 866 866 S -> T (IN REF. 1).

FT CONFLICT 1036 1036 A -> D (IN REF. 1).

FT CONFLICT 1074 1074 A -> G (IN REF. 1).

FT CONFLICT 1421 1421 A -> G (IN REF. 1).

FT CONFLICT 1545 1545 H -> T (IN REF. 1).

FT SEQUENCE 1694 AA; 185539 MW; C52427013F93178C CRC64;

Query Match 49.5%; Score 281.5; DB 1; Length 1694;

Best Local Similarity 48.2%; Pred. No. 1.1e-20;

Matches 54; Conservative 17; Mismatches 30; Indels 11; Gaps 2;

QY 1 LYKKNRYALKSGRLNAPNGVAENNDWPMGYTQEEARKNANNKNNRRIGDFGG 60

Db 601 LNLNYYIYALRGASTRSELPKNSGESNENWLYMGKTSDEAKRNVMNNHNNERNMGFNG 660

QY 61 FFDENKKGHNGALNLFNCKSAQNRLTGGANLNGG-----NCRP 102

Db 661 YFGECK-NNGNLNVTFKKGSEQNRLTGGTNGLDLKVKGKTLFLSGRP 711

RESULT 4
 IGA2_HAEIN STANDARD; PRT; 1702 AA.
 ID IGA2_HAEIN STANDARD; PRT; 1702 AA.
 AC P45384;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (EC 3.4.21.72) (IGAL PROTEASE).
 GN IGA.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OC NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HK393 / NCTC 8467 / SEROTYPE B;
 RX MEDLINE=92234949; PubMed=1373717;
 RA Poulsen K., Reinholdt J., Kilian M.;
 RT "A comparative genetic study of serologically distinct Haemophilus
 RT influenzae type 1 immunoglobulin A1 proteases.";
 RL J. Bacteriol. 174:2913-2921(1992).
 CC -1- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
 CC PRODUCING INTACT FC AND FAB FRAGMENTS.
 CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT
 CC CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE
 CC SUBSTRATES ARE KNOWN.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
 CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
 CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
 CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY

CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M87489; AAA24966.1; ;

DR MEROPS; S06.001; ;

DR InterPro; IPR000710; ;

DR PRINTS; PR00921; IGASERPTASE.

FW HydroLase; Serine protease; Transmembrane; Zymogen; Signal.

FT SIGNAL 1 25 POTENTIAL.

FT CHAIN 26 1014 IMMUNOGLOBULIN A1 PROTEASE.

FT PROPEP 1015 1702 HELPER PEPTIDE (POTENTIAL).

FT ACT_SITE 288 288 PROBABLE.

FT DOMAIN 1109 1124 2 X 8 AA TANDEM REPEATS OF A-K-V-E-K-E-E-

FT REPEAT 1109 1116 1.

FT REPEAT 1117 1124 2.

FT SEQUENCE 1702 AA; 186539 MW; 860F70D2667807A6 CRC64;

Query Match 49.5%; Score 281.5; DB 1; Length 1702;

Best Local Similarity 48.2%; Pred. No. 1.1e-20;

Matches 54; Conservative 17; Mismatches 30; Indels 11; Gaps 2;

QY 1 LYKKNRYALKSGRLNAPNGVAENNDWPMGYTQEEARKNANNKNNRRIGDFGG 60

Db 601 LNLNYYIYALRGASTRSELPKNSGESNENWLYMGKTSDEAKRNVMNNHNNERNMGFNG 660

QY 61 FFDENKKGHNGALNLFNCKSAQNRLTGGANLNGG-----NCRP 102

Db 661 YFGECK-NNGNLNVTFKKGSEQNRLTGGTNGLDLKVKGKTLFLSGRP 711

RESULT 5
 IGA3_HAEIN STANDARD; PRT; 1545 AA.
 ID IGA3_HAEIN STANDARD; PRT; 1545 AA.
 AC P45385;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE IMMUNOGLOBULIN A1 PROTEASE PRECURSOR (EC 3.4.21.72) (IGAL PROTEASE).
 GN IGA.
 OS Haemophilus influenzae.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Haemophilus.
 OC NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HK393 / NCTC 8467 / SEROTYPE B;
 RX MEDLINE=92234949; PubMed=1373717;
 RA Poulsen K., Reinholdt J., Kilian M.;
 RT "A comparative genetic study of serologically distinct Haemophilus
 RT influenzae type 1 immunoglobulin A1 proteases.";
 RL J. Bacteriol. 174:2913-2921(1992).
 CC -1- FUNCTION: VIRULENCE FACTOR; CLEAVES HOST IMMUNOGLOBULIN A
 CC PRODUCING INTACT FC AND FAB FRAGMENTS.
 CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF IMMUNOGLOBULIN A MOLECULES AT
 CC CERTAIN PRO-1-XAA BONDS IN THE HINGE REGION. NO SMALL MOLECULE
 CC SUBSTRATES ARE KNOWN.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- DOMAIN: THE SIGNAL PEPTIDE GUIDES THE PRECURSOR TO THE PERIPLASMIC
 CC SPACE, AND THE CARBOXY-TERMINAL HELPER DOMAIN ASSOCIATES WITH THE
 CC OUTER MEMBRANE TO FORM A PORE FOR EXCRETION OF THE PROTEASE
 CC DOMAIN. THE HELPER DOMAIN IS THEN RELEASED BY AUTOPROTEOLYSIS (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S6 (SERINE PROTEASE).


```

SQ SEQUENCE 1029 AA; 117081 MW; 99E2A1FC7CEB3D5A CRC64;

Query Match 12.7%; Score 72; DB 1; Length 1029;
Best Local Similarity 28.4%; Pred. No. 10;
Matches 25; Conservative 9; Mismatches 36; Indels 18; Gaps 4;

QY 25 GVAENNDWFMGYTQEQEARKNAMNKNRRIGDFGFFDEENGKNGALNLFN----- 79
   | : || : || : || : || : || : || : || : || : || : || : || : ||
Db 537 GFSNNN---KQYRQNNYNNNNNNNSNNGSYNNE---NNGNSYIKGNKNKFNKYYRP 590
   | : || : || : || : || : || : || : || : || : || : || : || : ||
QY 80 -----GKSAQNRELLTGGANLNGG-NG 100
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 591 SSSYTGKSPLSRYNMSYNHNNSSING 618

RESULT 12
POLSE_EV8V8 STANDARD; PRT; 1254 AA.
ID POLSE_EV8V8 STANDARD; PRT; 1254 AA.
AC P05674;
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE STRUCTURAL POLYPEPTIDE [CONTAINS: COAT PROTEIN C (EC 3.4.21.-); SPIKE
DE GLYCOPROTEINS E3, E2 AND E1; 6 KDA PEPTIDE].
OS Venezuelan equine encephalitis virus (strain TC-83).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCBI_TaxID=11037;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86306669; PubMed=3755750;
RA Johnson B.J.B., Kinney R.M., Kost C.L., Trent D.W.;
RT "Molecular determinants of alphavirus neurovirulence: nucleotide and
RT deduced protein sequence changes during attenuation of Venezuelan
RT equine encephalitis virus.";
RL J. Gen. Virol. 67:1951-1960(1986).
CC -|- FUNCTION: THE CAPSID PROTEIN IS AN AUTO-PROTEASE.
CC -|- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -|- MISCELLANEOUS: THE 6 KDA POLYPEPTIDE PROBABLY SERVES AS THE SIGNAL
CC SEQUENCE FOR THE MEMBRANE GLYCOPROTEIN E1, WHICH IS THE VIRAL
CC HEMAGGLUTININ.
CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X04368; CAA27883.1; -
CC PIR; A27871; VHWVVE.
CC HSP; P03315; IVCO.
CC MEROPS; S03.001; -
CC InterPro: IPR000930; -
CC InterPro: IPR000936; -
CC InterPro: IPR001836; -
CC InterPro: IPR002533; -
CC InterPro: IPR002548; -
CC Pfam; PF01589; Alpha_E1_glycop; 1.
CC Pfam; PF00943; Alpha_E2_glycop; 1.
CC Pfam; PF01563; Alpha_E3_glycop; 1.
CC Pfam; PF00944; Alpha_core; 1.
CC PRINTS; PR00798; TOGAVIRIN.
CC Coat protein; Polyprotein; Transmembrane; Glycoprotein; Hydrolase;
KW Serine protease.
FT CHAIN 1 275 COAT PROTEIN C (CAPSID PROTEIN C).
FT CHAIN 276 334 SPIKE GLYCOPROTEIN E3.
FT CHAIN 335 757 SPIKE GLYCOPROTEIN E2.
FT CHAIN 758 812 6 KDA PEPTIDE.
FT CHAIN 813 1254 SPIKE GLYCOPROTEIN E1.

```

```

FT ACT_SITE 152 152 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 158 158 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 226 226 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT TRANSMEM 702 718 POTENTIAL.
FT TRANSMEM 774 790 POTENTIAL.
FT TRANSMEM 792 808 POTENTIAL.
FT TRANSMEM 1231 1247 POTENTIAL.
FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 546 546 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 652 652 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1254 AA; 138485 MW; 7615698519A529F6 CRC64;

Query Match 12.6%; Score 71.5; DB 1; Length 1254;
Best Local Similarity 28.0%; Pred. No. 14;
Matches 28; Conservative 14; Mismatches 45; Indels 13; Gaps 5;

QY 9 YALKSGRLNAPDPENGVAENNDWFMGYTQEQEARKNAMNKN---NRRIGDFGFFDEE 65
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 139 YACVVGKGLFRPMHVEGKIDND--VLAALKTKKASKYDLEYADVDPQNNRADTFKYTHEKP 196
   ||| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY 66 NG--KGHNGALNLFNFKCKSAQNRELLTGGANLNGGNGRPV 103
   | : || : || : || : || : || : || : || : || : || : || : || : ||
Db 197 QGYSWHGHGAVQYE-NG-----RFTVPKGVGAKGDSGRPI 230
   | : || : || : || : || : || : || : || : || : || : || : || : ||

RESULT 13
POLSE_EV8V8 STANDARD; PRT; 1254 AA.
ID POLSE_EV8V8 STANDARD; PRT; 1254 AA.
AC P36330; Q66582; Q66583; Q66584; Q66585; Q66586;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE STRUCTURAL POLYPEPTIDE [CONTAINS: COAT PROTEIN C (EC 3.4.21.-); SPIKE
DE GLYCOPROTEINS E3, E2 AND E1; 6 KDA PEPTIDE].
OS Venezuelan equine encephalitis virus (strain Everglades Fe3-7c).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Togaviridae;
OC Alphavirus.
OX NCBI_TaxID=36383;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93187617; PubMed=8445371;
RA Snider J.M., Kinney R.M., Tsuchiya K.R., Trent D.W.;
RT "Molecular evidence that epizootic Venezuelan equine encephalitis
RT (VEE) I-AB viruses are not evolutionary derivatives of enzootic VEE
RT subtype I-E or II viruses.";
RL J. Gen. Virol. 74:519-523(1993).
CC -|- FUNCTION: THE CAPSID PROTEIN IS AN AUTO-PROTEASE.
CC -|- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -|- MISCELLANEOUS: THE 6 KDA POLYPEPTIDE PROBABLY SERVES AS THE SIGNAL
CC SEQUENCE FOR THE MEMBRANE GLYCOPROTEIN E1, WHICH IS THE VIRAL
CC HEMAGGLUTININ.
CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY S3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L04598; AAA42984.1; -
CC EMBL; L04598; AAA42983.1; ALT_TERM.
CC EMBL; L04598; AAA42985.1; ALT_SEQ.
CC EMBL; L04598; AAA42986.1; ALT_SEQ.
CC EMBL; L04598; AAA42987.1; ALT_SEQ.
CC EMBL; L04598; AAA42988.1; ALT_INIT.
CC PIR; JQ1978; JQ1978.
CC HSP; P03315; IVCO.
CC MEROPS; S03.001; -
CC InterPro: IPR000930; -

```


Search completed: May 1, 2001, 15:03:38
Job time: 285 sec

```

Query Match      12.6% Score 71.5; DB 1; Length 1255;
Best Local Similarity 28.0%; Pred. No. 14;
Matches 28; Conservative 14; Mismatches 45; Indels 13; Gaps 5;

QY 9 YALKSGRLNAPENGVAENNDWVFQYQTEEARKNAMNNKN---NRRIGDGFGGFDEE 65
   || || || || || || || || || || || || || || || || || || || ||
Db 139 YACVVGKGLFRPMHVEGKIDND--VLAALKTKKASYDLEADVQPNMRADTFYVTHEKP 196

QY 66 NG--KCHNGALNLFNCKSAQNRFLLTGGANLNGNGRVP 103
   | ||||| || || || || || || || || || || || || || || || || || ||
Db 197 QGYVSWHHGAVQYE-NG-----RFTVPKGVGAKGDSGRPI 230

```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 14:58:49 ; Search time 66.74 seconds
(without alignments)
89.077 Million cell updates/sec

Title: US-09-142-970-4

Perfect score: 569

Sequence: 1 LYKNRYALKSGRLNAP.....NRLLTGGANLGGNGRPVK 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 390729 seqs, 57163235 residues

Total number of hits satisfying chosen parameters: 390729

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0401.*
1: /SID56/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID56/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID56/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID56/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID56/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID56/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID56/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID56/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID56/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID56/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID56/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID56/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID56/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID56/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID56/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID56/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID56/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID56/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID56/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID56/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID56/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID56/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	569	100.0	104	19 W61605	Neisseria IgA1 pro
2	524	92.1	104	19 W61606	Neisseria IgA1 pro
3	475	83.5	104	19 W61603	Neisseria IgA1 pro
4	462	81.2	104	19 W61604	Neisseria IgA1 pro
5	453	79.6	105	19 W65656	105-mer peptide us
6	453	79.6	105	19 W61602	Neisseria IgA1 pro
7	281.5	49.5	1541	11 R07304	IgA1 protease. Ha
8	134	23.6	1431	21 R75564	Neisseria meningit
9	107.5	18.9	1468	20 Y38825	Neisseria gonorrh
10	91	16.0	1394	17 R32768	Adhesion and penet
11	90.5	15.9	1454	21 Y56621	Neisseria meningit

12	90.5	15.9	1457	20 Y38823	Neisseria meningit
13	90.5	15.9	1457	21 B58592	N. meningitidis am
14	90.5	15.9	1457	21 B25662	N. meningitidis am
15	90.5	15.9	1457	21 Y56622	Neisseria meningit
16	83.5	14.7	1449	20 Y38824	Neisseria meningit
17	80	14.1	461	16 R75506	Nicotiana alata ar
18	78.5	13.8	807	21 B18311	Plasmodium falcipa
19	76	13.4	1577	17 R91047	Alpha-D-glucosyltr
20	74.5	13.1	631	16 R77896	Bacterial transfer
21	74.5	13.1	631	18 W08970	Amino acid sequenc
22	74.5	13.1	631	19 W54128	H. influenzae stra
23	74.5	13.1	631	21 Y51784	H. influenzae non-
24	74.5	13.1	631	21 Y80366	H. influenzae stra
25	73	12.8	115	21 G04598	Arabidopsis thalia
26	73	12.8	537	7 P60452	Sequence of the As
27	70.5	12.4	281	21 Y90256	Streptococcus equi
28	70	12.3	764	21 B18282	Plasmodium falcipa
29	70	12.3	1817	21 B18255	Plasmodium falcipa
30	69.5	12.2	263	21 G42442	Arabidopsis thalia
31	69.5	12.2	263	21 G57383	Arabidopsis thalia
32	69.5	12.2	263	21 G59074	Arabidopsis thalia
33	69.5	12.2	280	21 B63197	Gene 21 human secr
34	69.5	12.2	287	21 G42441	Arabidopsis thalia
35	69.5	12.2	288	21 G57382	Arabidopsis thalia
36	69.5	12.2	447	17 W01622	Oyster pearl prote
37	69.5	12.2	954	21 B48550	Ruminococcus flave
38	69	12.1	957	21 Y68674	Amino acid sequenc
39	68.5	12.0	597	12 R12384	Bovine pancreatic
40	68.5	12.0	631	16 R77894	Bacterial transfer
41	68	12.0	971	21 Y95687	Cosmid CHRIM5 enco
42	68	12.0	1278	21 B18277	Plasmodium falcipa
43	67.5	11.9	437	16 R75507	Nicotiana plumbagi
44	67.5	11.9	441	21 Y75917	Neisseria strain z
45	67.5	11.9	441	21 Y75918	Neisseria strain z

ALIGNMENTS

RESULT 1
W61605
ID W61605 standard; peptide; 104 AA.
XX
AC W61605;
XX
XX
DT 27-OCT-1998 (first entry)
XX
DE Neisseria IgA1 protease fragment 4.
XX
KW Immunoglobulin protease; carrier; paediatric; vaccine;
KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.
XX
OS Neisseria sp.
XX
PN W09831791-Al.
XX
PD 23-JUL-1998.
XX
PF 20-JAN-1998; 98WO-EP00294.
XX
PR 21-JAN-1997; 97EP-0100883.
XX
XX
(PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
(INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.
XX
PI Achtmann M, Moreau M;
XX
XX 105-mer peptide us
DR WPI; 1998-414092/35.
XX
XX New peptide from Neisseria immunoglobulin protease - useful as
PT immunogenic carrier, e.g. particularly for polysaccharides),
PT forming conjugates used in vaccines against Neisseria and
PT Haemophilus

RESULT 6

W61602
ID W61602 standard; peptide; 105 AA.

XX AC W61602;

XX DT 27-OCT-1998 (first entry)

XX DE Neisseria IgA1 protease fragment 1.

XX KW Immunoglobulin protease; carrier; paediatric; vaccine;
XX KW epidemic bacterial infection; Neisseria; Haemophilus; T-cell.

XX OS Neisseria sp.

XX PN W09831791-A1.

XX PD 23-JUL-1998.

XX PF 20-JAN-1998; 98WO-EP00294.

XX PR 21-JAN-1997; 97EP-0100883.

XX PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX PA (INMR) PASTEUR MERIEUX SERUMS & VACCINS SA.

XX PI Achtmann M, Moreau M;

XX XX WPI; 1998-414092/35.

XX XX New peptide from Neisseria immunoglobulin protease - useful as
XX PT immunogenic carrier, e.g. particularly for polysaccharide(s),
XX PT forming conjugates used in vaccines against Neisseria and
XX PT Haemophilus

XX PS Claim 6; Page 10; 32pp; English.

XX XX The Neisseria immunoglobulin protease fragments W61602-W61606 are used
XX CC as carriers for a conjugate, particularly in combination with a
XX CC polysaccharide. They can be used in paediatric or other vaccines,
XX CC particularly for prevention of epidemic bacterial infections, especially
XX CC those caused by Neisseria or Haemophilus. The protease fragment is a
XX CC highly immunogenic carrier that elicits a T-cell response, resulting in a
XX CC long-lasting memory and high antibody titre, and possibly making possible
XX CC vaccination without adjuvant.

XX SO Sequence 105 AA;

Query Match 79.6%; Score 453; DB 19; Length 105;
Best Local Similarity 84.5%; Pred. No. 1.7e-44;
Matches 82; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LYKNRYRYALKSGRLNAPMPENGVAENNDWFMGYTQEEARKNAMNKNRRIGDFGG 60

Db 2 lyknryryalksgsgvnapmpengqtenndwilmgstqeaknmmhknqrisgfsq 61

Qy 61 FDEENGKGHGHALNLFNFKSAQNRFLLTGGANLNG 97

Db 62 ffggeegkgghginalnfnfgksaqnrflttgtnlng 98

RESULT 7

R07304
ID R07304 standard; protein; 1541 AA.

XX AC R07304;

XX DT 31-JAN-1991 (first entry)

XX DE IgA1 protease.

XX XX IgA1; vaccine; meningitis; gonorrhoea; allergies.
XX OS Haemophilus influenzae.

XX PN W09011367-A.

XX PD 04-OCT-1990.

XX PF 16-MAR-1990; 90WO-DK00073.

XX PR 17-MAR-1989; 89DK-0001308.

XX PA (KILI/) KILIAN M.

XX PI Kilian M, Poulsen K;

XX DR WPI; 1990-320267/42.

XX DR N-PSDB; Q06164.

XX PT Immunoglobulin A1 protease prodn. - by cloning from
XX PT microorganisms for immunisation against immunoglobulin A1
XX PT protease producing bacteria

XX PS Disclosure; fig 3; 44pp; English.

XX CC This immunoglobulin (Ig)A1 protease is produced by recombinant DNA
XX CC methods. It is useful in a vaccine for e.g. meningococcal meningitis,
XX CC gonorrhoea or allergic diseases. It specifically cleaves the heavy
XX CC chain of human IgA1 in the hinge region.

XX SQ Sequence 1541 AA;

Query Match 49.5%; Score 281.5; DB 11; Length 1541;
Best Local Similarity 48.2%; Pred. No. 2.1e-23;
Matches 54; Conservative 17; Mismatches 30; Indels 11; Gaps 2;

Qy 1 LYKNRYRYALKSGRLNAPMPENGVAENNDWFMGYTQEEARKNAMNKNRRIGDFGG 60

Db 595 Inlenytyairkgastrselpkngsgesnewlymgktsdeaknvmhlnnermngfng 654

Qy 61 FDEENGKGHGHALNLFNFKSAQNRFLLTGGANLNGG-----NGRP 102

Db 655 yfgeegkg-nngnlnvtfkqseqnrflttgtnlngdltvekgflflsgrp 705

RESULT 8

Y75564
ID Y75564 standard; Protein; 1431 AA.

XX AC Y75564;

XX DT 21-MAR-2000 (first entry)

XX DE Neisseria meningitidis ORF 759 protein sequence SEQ ID NO:2602.

XX KW Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
XX KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
XX KW antibacterial; gene therapy.

XX OS Neisseria meningitidis.

XX PN W09957280-A2.

XX PD 11-NOV-1999.

XX PF 30-APR-1999; 99WO-US09346.

XX PR 01-MAY-1996; 98US-0083758.

XX PR 31-JUL-1998; 98US-0094869.

XX PR 02-SEP-1998; 98US-0098994.

XX PR 02-SEP-1998; 98US-0099062.

Se 8 10 NOV 04

FT Protein 26...959
 FT /note= "Secreted 110-kDa protease fragment"
 FT Domain 241...248
 FT /note= "Putative catalytic domain"
 FT Active-site 243
 FT /note= "Active site serine residue"
 FT Region 698
 FT /note= "Conserved cysteine residue"
 FT Region 709
 FT /note= "Conserved cysteine residue"
 FT Cleavage-site 960
 FT /note= "Site for cleavage of secreted protease from
 FT outer membrane protein"
 FT Protein 960...1394
 FT /note= "45-kDa outer membrane protein fragment"
 FT
 PN WO9605858-A1.
 XX
 XX 29-FEB-1996.
 XX
 XX 16-AUG-1995; 95WO-US10661.
 XX
 XX 25-AUG-1994; 94US-0296791.
 XX
 XX (STRD) UNIV LELAND STANFORD JUNIOR.
 XX (UNIW) UNIV WASHINGTON.
 XX
 XX Falkow S, St Geme JW;
 XX WPI; 1996-151147/15.
 DR N-PSDB; T17215.
 XX
 XX Haemophilus adhesion and penetration protein and corresponding DNA
 PT - used to produce vaccines against H. influenzae infection
 PT
 PS Claim 2; Fig 6; 105pp; English.
 XX
 CC The sequence represents a Haemophilus influenzae adhesion and
 CC penetration protein, encoded by the hap gene. The protein is
 CC first synthesised as a preprotein with 3 functional domains (the
 CC N-terminal signal peptide, the protease and a C-terminal helper
 CC domain), which is transported to the periplasm, followed by
 CC insertion of the C-terminal beta-domain into the outer membrane,
 CC possibly forming a pore, and export of the N-terminal fragment
 CC through the outer membrane, followed by autolytic cleavage
 CC and secretion of the mature protease, leaving an outer membrane
 CC protein fragment. An active site serine residue characteristic of
 CC serine proteases is present, along with 2 cysteine residues which
 CC may be important in forming the catalytic domain. A gene encoding
 CC the protein may be inserted in a vector and expressed in
 CC recombinant host cells, for use as a recombinant vaccine. The
 CC protein may also be used in diagnostic monoclonal antibody production.
 XX
 SQ Sequence 1394 AA;

Query Match 16.0%; Score 91; DB 17; Length 1394;
 Best Local Similarity 32.4%; Pred. No. 0.12;
 Matches 22; Conservative 10; Mismatches 26; Indels 10; Gaps 1;

QY 45 NAMNNKRRGDFGFFDEENGKNGALNLFNGKSAQNRFLTGTGANLNGG----- 98
 Db 551 nnkldyrkeiayngwfgedknkngnllykpttedrlllsggtlnkgdtkg 610
 QY 99 ----NGRP 102
 Db 611 klffsgrp 618

RESULT 11
 Y56621
 ID *Y56621--standard; Protein; 1454 AA.
 XX

AC Y56621;
 XX
 DT 21-FEB-2000 (first entry)
 XX
 DE Neisseria meningitidis strain ATCC13090 BASB006 protein sequence.
 XX
 KW Neisseria meningitidis; BASB006; diagnosis; bacterial; infection;
 KW vaccine; antibiotic; upper respiratory tract infection; meningitis;
 KW invasive bacterial disease; bacteraemia; screening; antibacterial.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO955873-A2.
 XX
 PD 04-NOV-1999.
 XX
 PF 20-APR-1999; 99WO-EP02766.
 XX
 PR 24-APR-1998; 98GB-0008866.
 XX
 XX (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 PI Thonnard J;
 XX
 XX WPI; 2000-052810/04.
 DR N-PSDB; Z38917.
 XX
 XX Novel polynucleotides and polypeptides from Neisseria meningitis used
 PT to prepare vaccines against bacterial infections
 PT
 XX
 PS Claim 3; Page 86-91; 103pp; English.
 XX
 CC The present sequence is BASB006 isolated from Neisseria meningitidis
 CC strain ATCC13090. BASB006 polynucleotides and polypeptides may be
 CC employed as research reagents and material for the discovery of
 CC treatments and diagnostics for diseases, particularly human diseases.
 CC They can be used for diagnosis of disease, staging of disease, or
 CC determining response of an infectious organism to drugs. The
 CC polynucleotides may be used as a source for hybridisation probes, and for
 CC screening of genetic mutations, serotype, organism or strain
 CC identification, identification of mutation in BASB013 sequences, and as
 CC components of arrays which are useful for diagnostic and prognostic
 CC purposes. The polypeptides can be used to produce antibodies. The
 CC polypeptides can also be used in vaccine formulations, and to identify
 CC agonists and antagonists. The polypeptides, antibodies, agonists and
 CC antagonists (which are bacteriostatic) are used for the treatment and
 CC prevention of diseases such as upper respiratory tract infection,
 CC invasive bacterial diseases such as bacteraemia and meningitis, and for
 CC the development and screening of antibacterial drugs. They are also used
 CC in the prevention of adhesion of bacteria to eukaryotic matrix proteins
 CC on in-dwelling devices, or to extracellular proteins on wounds, and to
 CC prevent tissue damage and/or block the normal progression of pathogenesis
 CC in infections initiated other than by the implantation of in-dwelling
 CC devices or by other surgical techniques.
 XX
 SQ Sequence 1454 AA;

Query Match 15.9%; Score 90.5; DB 21; Length 1454;
 Best Local Similarity 25.6%; Pred. No. 0.14;
 Matches 32; Conservative 20; Mismatches 38; Indels 35; Gaps 7;

QY 8 YKALSGGRRLNMPMPENGVA-----ENNND--WVPMGYTQEE-----ARKNANNKNN 52
 Db 527 yfgfr-ggrldl----nghslsfhrlqntdegamivnnhqdkestvtitgnkdattgnn 581
 QY 53 RRGD-----FGFFDEENGKNGALNLFNGKSAQNRFLTGTGANLNGG----- 98
 Db 582 nsldskkeiayngwfgedtktngrlnlvyqpaaedrtlllsggtlnngtngtkl 641
 QY 99 -NGRP 102
 Db 642 fsgrp 646

Sequence .1457 AA;

CC on in-dwelling devices, or to extracellular proteins on wounds, and to
CC prevent tissue damage and/or block the normal progression of pathogenesis
CC in infections initiated other than by the implantation of in-dwelling
CC devices or by other surgical techniques.

[illegible]

Search completed: May 1, 2001, 14:58:50
Job time: 612 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 14:57:18 ; Search time 149.44 Seconds
(without alignments)
111.944 Million cell updates/sec

Title: US-09-142-970-1

Perfect score: 558

Sequence: 1 LYXKNRYRYALKSGGSVNAP.....NRFLITGTGTLNKGKISVTQG 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

- 1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
- 2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
- 3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
- 4: /cgn2_6/ptodata/2/paa/US080_COMB.pep.*
- 5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
- 6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
- 7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
- 8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
- 9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
- 10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
- 11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
- 12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
- 13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
- 14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
- 15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
- 16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
- 17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
- 18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
- 19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
- 20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
- 21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
- 22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
- 23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	558	100.0	104	15	US-09-142-970-1
2	558	100.0	105	15	US-09-155-077-1
3	558	100.0	1560	9	US-08-578-788-4
4	540	96.8	104	15	US-09-142-970-2
5	537	96.2	104	15	US-09-142-970-3
6	494	88.5	104	15	US-09-142-970-5
7	453	81.2	104	15	US-09-142-970-4
8	329.5	59.1	1540	9	US-08-578-788-2
9	329.5	59.1	1541	6	US-08-296-791-3
10	322.5	57.8	1702	6	US-08-296-791-5

11	320.5	57.4	1545	6	US-08-296-791-4	Sequence 4, Appli
12	311.5	55.8	1848	6	US-08-296-791-6	Sequence 6, Appli
13	114	20.4	1394	6	US-08-296-791-2	Sequence 2, Appli
14	84	15.1	1201	23	US-60-167-217-7969	Sequence 7969, Ap
15	84	15.1	1201	23	US-60-173-464-6367	Sequence 6367, Ap
16	84	15.1	1225	23	US-60-191-637-7869	Sequence 7869, Ap
17	84	15.1	1225	23	US-60-191-681-6135	Sequence 6135, Ap
18	83	14.9	450	21	US-09-733-089-13711	Sequence 13711, A
19	78	14.0	658	13	US-08-993-002A-9518	Sequence 9518, Ap
20	78	14.0	996	11	US-08-761-066-544	Sequence 544, App
21	78	14.0	996	13	US-08-993-002A-9519	Sequence 9519, Ap
22	75.5	13.6	229	16	US-09-248-796-24831	Sequence 24831, A
23	75.5	13.5	280	1	PCT-US00-09066-123	Sequence 123, App
24	75.5	13.5	954	19	US-09-570-856-41	Sequence 41, Appli
25	74.5	13.4	705	9	US-08-591-447-6	Sequence 6, Appli
26	74.5	13.4	705	9	US-08-591-447A-6	Sequence 6, Appli
27	74.5	13.4	1004	23	US-60-167-217-17901	Sequence 17901, A
28	74.5	13.4	1004	23	US-60-171-627-1416	Sequence 1416, Ap
29	74.5	13.4	1004	23	US-60-173-464-14681	Sequence 14681, A
30	74.5	13.4	1366	23	US-60-191-637-27784	Sequence 27784, A
31	74.5	13.4	1366	23	US-60-191-681-22393	Sequence 22393, A
32	74.5	13.4	1594	23	US-60-191-637-17909	Sequence 17909, A
33	74.5	13.4	1594	23	US-60-191-681-14162	Sequence 14162, A
34	74	13.3	1838	23	US-60-191-637-19526	Sequence 19526, A
35	74	13.3	1838	23	US-60-191-681-15404	Sequence 15404, A
36	74	13.3	1864	23	US-60-167-217-19606	Sequence 19606, A
37	74	13.3	1864	23	US-60-173-464-16004	Sequence 16004, A
38	73.5	13.2	631	8	US-08-482-811-115	Sequence 115, App
39	73.5	13.2	631	8	US-08-483-577-115	Sequence 115, App
40	73.5	13.2	631	10	US-08-637-654-115	Sequence 115, App
41	73.5	13.2	631	10	US-08-649-518-115	Sequence 115, App
42	73.5	13.2	631	12	US-08-897-438-115	Sequence 115, App
43	73.5	13.2	631	13	US-08-973-721A-115	Sequence 115, App
44	73	13.1	329	6	US-08-204-808-58	Sequence 58, Appli
45	73	13.1	329	8	US-08-461-651-58	Sequence 58, Appli

ALIGNMENTS

RESULT 1
US-09-142-970-1
; Sequence 1, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtmann
; APPLICANT: Monique Moreau
; TITLE OF INVENTION: IGAI PROTEASE FRAGMENT AS CARRIER
; FILE REFERENCE: 7101/0EG16US0
; CURRENT APPLICATION NUMBER: US/09/142,970
; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-142-970-1

Query Match 100.0%; Score 558; DB 15; Length 104;
Best Local Similarity 100.0%; Pred. No. 1.5e-59;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LYXKNRYRYALKSGGSVNAPMPENGQTNNDWILMGSTQEEAKKNMHNKNNRISGFSG 60
DB 1 LYXKNRYRYALKSGGSVNAPMPENGQTNNDWILMGSTQEEAKKNMHNKNNRISGFSG 60
QY 61 FFGEENKKGHNGALNLFNFKSAQNRELLTGTGTLNKGKISVTQG 104

RESULT

OTHER
OTHER
OTHER


```
Query Match 10.6%; Score 59; DB 5; Length 498;
Best Local Similarity 31.3%; Pred. No. 7.3;
Matches 21; Conservative 10; Mismatches 16; Indels 20; Gaps 4;

QY 48 NHKNNOR--ISGSGPGEENGKNGHNCALNLFNGKSAQN-----RFLTTGG----- 92
Db 396 NEQNYRLYLKGTGTAGROSSLILHGA---DFSTKADNDNCMKCALMLTGGWFFDAC 452
QY 93 --TNLNG 97
Db 453 GPSNLNG 459

RESULT 7
US-09-739-449-9987
; Sequence 9987, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 9987
; LENGTH: 1976
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9987

Query Match 10.0%; Score 56; DB 5; Length 1976;
Best Local Similarity 24.7%; Pred. No. 91;
Matches 23; Conservative 10; Mismatches 36; Indels 24; Gaps 3;

QY 1 LYXKNRYRYALKSGGSVNAPMPENGQTNNDWILMGSTOEAKKNAMHNKNNORISGFSG 60
Db 1281 LTFDNLTVDLKSGLEVNRIVGTNGAD-----VLVGTA---AAETMFGLEGNDELQGFSG 1332
QY 61 FFGENGKGHNGALNLFNGKSAQNRLTGGT 93
Db 1333 DRLDGGAGDDA-----LAGGT 1349

RESULT 8
PCT-US01-08117-19
; Sequence 19, Application PC/TUS0108117
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: MATHUR, Preete
; APPLICANT: SHAH, Purvi
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: REDDY, Roopa
; TITLE OF INVENTION: TRANSCRIPTION FACTORS
; FILE REFERENCE: PF-0761 PCT
; CURRENT APPLICATION NUMBER: PCT/US01/08117
; CURRENT FILING DATE: 2001-03-14
; NUMBER OF SEQ ID NOS: 214
; SOFTWARE: PERL Program
; SEQ ID NO 19
; LENGTH: 549
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: Incyte ID No: 1238747CDI
PCT-US01-08117-19

Query Match 9.9%; Score 55.5; DB 1; Length 549;
Best Local Similarity 31.7%; Pred. No. 21;
Matches 19; Conservative 6; Mismatches 24; Indels 11; Gaps 3;

QY 1 LYXKNRYRYALKSGGSVNAPMPENGQTNNDWILMGSTOEAKKNAMHNKNNORISGFSG 59
Db 475 LYKKK-----MKHGSSVKCIRNEDG-----TWLTPNEFEVEGKGRNAKNNKRNIRCEGWT 524

RESULT 9
US-09-792-024-105
; Sequence 105, Application US/09792024
; GENERAL INFORMATION:
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug
; TITLE OF INVENTION: Targets Discovery
; FILE REFERENCE: 10182-004-999
; CURRENT APPLICATION NUMBER: US/09/792,024
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 490
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 768
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-792-024-105

Query Match 9.9%; Score 55.5; DB 5; Length 768;
Best Local Similarity 26.2%; Pred. No. 32;
Matches 27; Conservative 12; Mismatches 35; Indels 29; Gaps 5;

QY 15 GSVNAPM--PENGQTNNDWILMGSTQ-----LTGPTKKRKSIAKLEAAELAKESDTNDKQETSESEE 60
Db 334 GDINSFLPKKNGQ-----LTGPTKKRKSIAKLEAAELAKESDTNDKQETSESEE 386
QY 61 FFGENGKGHNGALNLFNGKSAQNRLTGGTNLNGKISVTQ 103
Db 387 --GEADAGHSDVSN-----SPVERILELGGEGGNTSLILEQ 421

RESULT 10
PCT-US01-08117-24
; Sequence 24, Application PC/TUS0108117
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: PATTERSON, Chandra
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: MATHUR, Preete
; APPLICANT: SHAH, Purvi
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: REDDY, Roopa
; TITLE OF INVENTION: TRANSCRIPTION FACTORS
; FILE REFERENCE: PF-0761 PCT
; CURRENT APPLICATION NUMBER: PCT/US01/08117
; CURRENT FILING DATE: 2001-03-14
```

OTHER INFORMATION: Incyte ID No: 1377380CD1
PCT-US01-08117-24

Qy 66 NGKGHNGALNLFNGKSAQNRFLLTGGTNLNGKISVT 102
 ||:|: : : ||:| : || : ||:| :
 Db 81 NGRGVHAGILKSMGKTAQ----ILGG-NOSGRVOVS 112

RESULT 13
US-09-420-785A-4
: Sequence 4, Application US/09420785A

```

; FASTA ID: 200004_24
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30
; LENGTH: 287
; TYPE: PrT
; ORGANISM: Homo sapiens
PCT-US01-11797-30

```

RESULT 14
US-09-661-322A-63
; Sequence 63, Application US/09661322A

: GENERAL INFORMATION:
 : APPLICANT: Baum, James A.
 : APPLICANT: Chu, Chih-Rei
 : APPLICANT: Donovan, William P.
 : APPLICANT: Gilmer, Mark J.
 : APPLICANT: Rupert, Mark J.
 : TITLE OF INVENTION: Lepidopteran-Active Ba
 : TITLE OF INVENTION: and Methods of Use

Query Match 9.6%; Score 53.5; DB 5; Length 1227;
Best Local Similarity 27.6%; Pred. No. 98;

Matches 27; Conservative 13; Mismatches 35; Indels 23; Gaps 5;
QY 5 NYRYA-----LKGSGYNAPM-----PENGQTEENNNDWILMGSTQEEAKKNAMNHK 50
Db 461 NYESYSHRLSNIRLISGNTLRAPVYKSWTHRSADRTNTISSDSI-----TQPLVKSFNLNS 516
QY 51 NNORISGSGFFGEE-----NGKGHNGALNLFNCKSAQ 84
Db 517 GTSVVG-PGFTGGDIIRTNVNGSVLSMGLNFNNTSLQ 553
RESULT 15
US-09-739-449-8869
; Sequence 8869, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 8869
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-8869

Query Match 9.5%; Score 53; DB 5; Length 273;
Best Local Similarity 22.2%; Pred. No. 17;
Matches 24; Conservative 12; Mismatches 32; Indels 40; Gaps 4;
QY 5 NYRYAALKSGSGYNAPMPENGQTEENNNDWILMGSTQEEAKKNAMNHKNQRIQSF-----SG 60
Db 81 NNRHGAELITAAITLAAHAAKGGSLGPDVILFGGDD-----HASRERIDGFHAAKAD 132
QY 61 FFGGNGK-----GHNGAL-----NLNFNG 80
Db 133 YFGVEGGDDIEITGYSPHMTMAFERFFGRRRLPRCFFVNSSINFEG 180

Search completed: May 1, 2001, 15:03:03
Job time: 319 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 15:02:49 ; Search time 149.44 Seconds
(without alignments)
111.944 Million cell updates/sec

Title: US-09-142-970-2

Perfect score: 562

Sequence: 1 LYKNRYRYALKSGGSVNAP.....NRFLTGTGTLNGLKISVTQG 104

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1009251 seqs, 160854530 residues

Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*

2: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*

3: /cgn2_6/ptodata/2/paa/US08_COMB.pep.*

4: /cgn2_6/ptodata/2/paa/US09_COMB.pep.*

5: /cgn2_6/ptodata/2/paa/US10_COMB.pep.*

6: /cgn2_6/ptodata/2/paa/US11_COMB.pep.*

7: /cgn2_6/ptodata/2/paa/US12_COMB.pep.*

8: /cgn2_6/ptodata/2/paa/US13_COMB.pep.*

9: /cgn2_6/ptodata/2/paa/US14_COMB.pep.*

10: /cgn2_6/ptodata/2/paa/US15_COMB.pep.*

11: /cgn2_6/ptodata/2/paa/US16_COMB.pep.*

12: /cgn2_6/ptodata/2/paa/US17_COMB.pep.*

13: /cgn2_6/ptodata/2/paa/US18_COMB.pep.*

14: /cgn2_6/ptodata/2/paa/US19_COMB.pep.*

15: /cgn2_6/ptodata/2/paa/US20_COMB.pep.*

16: /cgn2_6/ptodata/2/paa/US21_COMB.pep.*

17: /cgn2_6/ptodata/2/paa/US22_COMB.pep.*

18: /cgn2_6/ptodata/2/paa/US23_COMB.pep.*

19: /cgn2_6/ptodata/2/paa/US24_COMB.pep.*

20: /cgn2_6/ptodata/2/paa/US25_COMB.pep.*

21: /cgn2_6/ptodata/2/paa/US26_COMB.pep.*

22: /cgn2_6/ptodata/2/paa/US27_COMB.pep.*

23: /cgn2_6/ptodata/2/paa/US28_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	562	100.0	104	15	US-09-142-970-2
2	546	97.2	104	15	US-09-142-970-3
3	540	96.1	104	15	US-09-142-970-1
4	540	96.1	103	15	US-09-155-077-1
5	540	96.1	1560	9	US-08-578-788-4
6	514	91.5	104	15	US-09-142-970-5
7	475	84.5	104	15	US-09-142-970-4
8	329.5	58.6	1540	9	US-08-578-788-2
9	329.5	58.6	1541	6	US-08-296-791-3
10	322.5	57.4	1702	6	US-08-296-791-5

11	319.5	56.9	1545	6	US-08-296-791-4	Sequence 4, Appli
12	312.5	55.6	1848	6	US-08-296-791-6	Sequence 6, Appli
13	114	20.3	1394	6	US-08-296-791-2	Sequence 2, Appli
14	89	15.8	450	21	US-09-733-089-13711	Sequence 13711, A
15	82.5	14.7	426	23	US-60-161-932-2066	Sequence 2066, Ap
16	82.5	14.7	464	23	US-60-167-217-11345	Sequence 11345, A
17	82.5	14.7	464	23	US-60-167-217-11363	Sequence 11363, A
18	82.5	14.7	464	23	US-60-173-464-9184	Sequence 9184, Ap
19	82.5	14.7	464	23	US-60-173-464-9196	Sequence 9196, Ap
20	82.5	14.7	464	23	US-60-191-637-11351	Sequence 11351, A
21	82.5	14.7	464	23	US-60-191-637-11369	Sequence 11369, A
22	82.5	14.7	464	23	US-60-191-681-8894	Sequence 8894, Ap
23	82.5	14.7	464	23	US-60-191-681-8906	Sequence 8906, Ap
24	81	14.4	1201	23	US-60-167-217-7969	Sequence 7969, Ap
25	81	14.4	1201	23	US-60-173-464-6367	Sequence 6367, Ap
26	81	14.4	1225	23	US-60-191-637-7869	Sequence 7869, Ap
27	81	14.4	1225	23	US-60-191-681-6135	Sequence 6135, Ap
28	79.5	14.1	705	9	US-08-591-447-6	Sequence 6, Appli
29	79.5	14.1	705	9	US-08-591-447A-6	Sequence 6, Appli
30	79	14.1	419	16	US-09-270-767-41767	Sequence 41767, A
31	79	14.1	586	23	US-60-167-217-15320	Sequence 15320, A
32	79	14.1	586	23	US-60-171-625-324	Sequence 324, App
33	79	14.1	586	23	US-60-173-464-12542	Sequence 12542, A
34	79	14.1	586	23	US-60-191-637-15302	Sequence 15302, A
35	79	14.1	586	23	US-60-191-637-28261	Sequence 28261, A
36	79	14.1	586	23	US-60-191-681-12116	Sequence 12116, A
37	79	14.1	586	23	US-60-191-681-22870	Sequence 22870, A
38	79	14.1	586	23	US-60-219-005-174	Sequence 174, App
39	79	14.1	586	23	US-60-219-005-297	Sequence 297, App
40	78.5	14.0	1073	15	US-09-167-568-45	Sequence 45, Appl
41	78.5	14.0	1073	15	US-09-167-568A-45	Sequence 45, Appl
42	78.5	14.0	1073	16	US-09-206-942-49	Sequence 49, Appl
43	78.5	14.0	1079	15	US-09-167-568-43	Sequence 43, Appl
44	78.5	14.0	1079	15	US-09-167-568A-43	Sequence 43, Appl
45	78.5	14.0	1079	16	US-09-206-942-47	Sequence 47, Appl

ALIGNMENTS

RESULT 1

US-09-142-970-2

; Sequence 2, Application US/09142970

; GENERAL INFORMATION:

; APPLICANT: Mark Achtmann

; APPLICANT: Monique Moreau

; TITLE OF INVENTION: IGAL PROTEASE FRAGMENT AS CARRIER

; TITLE OF INVENTION: PEPTIDE

; FILE REFERENCE: 7101/0EG16USO

; CURRENT APPLICATION NUMBER: US/09/142,970

; CURRENT FILING DATE: 1999-04-02

; EARLIER APPLICATION NUMBER: EP 97100883.4

; EARLIER FILING DATE: 1997-01-21

; EARLIER APPLICATION NUMBER: PCT/EP98/00294

; EARLIER FILING DATE: 1998-01-20

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 104

; TYPE: PRT.

; ORGANISM: Neisseria meningitidis

US-09-142-970-2

Query Match 100.0%; Score 562; DB 15; Length 104;

Best Local Similarity 100.0%; Pred. No. 4.1e-61;

Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LYKNRYRYALKSGGSVNAPENGVTENNDDWFMVGTQBEAKKNAMHNKNNORISGFSG 60

|||||

Db 1 LYKNRYRYALKSGGSVNAPENGVTENNDDWFMVGTQBEAKKNAMHNKNNORISGFSG 60

QY 61 FFGENGKGHNGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 104

Db 61 PFGEENGKNGALNLFNGKSAQNRFLLTGCTNLNGKISVTQG 104
|||||

RESULT 2

US-09-142-970-3
; Sequence 3, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtman
; APPLICANT: Monique Moreau
; TITLE OF INVENTION: IGAL PROTEASE FRAGMENT AS CARRIER
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE: 7101/0E616U50
; CURRENT APPLICATION NUMBER: US/09/142,970
; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-142-970-3

Query Match 97.2%; Score 546; DB 15; Length 104;
Best Local Similarity 97.1%; Pred. No. 3.7e-59;
Matches 101; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 LYXKNRYALKSGSVNAPMPGVNTENNDDWVFMGYTQEEAKKNAMHKNQRIISGFSG 60
Db 1 LYXKNRYALKSGSVNAPMPGVNTENNDDWVFMGYTQEEAKKNAMHKNQRIISGFSG 60
Qy 61 PFGEENGKNGALNLFNGKSAQNRFLLTGCTNLNGKISVTQG 104
Db 61 PFGEENGKNGALNLFNGKSAQNRFLLTGCTNLNGKISVTQG 104

RESULT 3

US-09-142-970-1
; Sequence 1, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtman
; APPLICANT: Monique Moreau
; TITLE OF INVENTION: IGAL PROTEASE FRAGMENT AS CARRIER
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE: 7101/0E616U50
; CURRENT APPLICATION NUMBER: US/09/142,970
; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-142-970-1

Query Match 96.1%; Score 540; DB 15; Length 104;
Best Local Similarity 96.2%; Pred. No. 2e-58;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 1 LYXKNRYALKSGSVNAPMPGVNTENNDDWVFMGYTQEEAKKNAMHKNQRIISGFSG 60
Db 1 LYXKNRYALKSGSVNAPMPGVNTENNDDWVFMGYTQEEAKKNAMHKNQRIISGFSG 60

Qy 61 PFGEENGKNGALNLFNGKSAQNRFLLTGCTNLNGKISVTQG 104
Db 61 PFGEENGKNGALNLFNGKSAQNRFLLTGCTNLNGKISVTQG 104

RESULT 4

US-09-155-077-1
; Sequence 1, Application US/09155077
; GENERAL INFORMATION:
; APPLICANT: Moreau Monique
; APPLICANT: Mistretta, Noelle
; TITLE OF INVENTION: Polysaccharide-Peptide-Conjugates
; FILE REFERENCE: MBHB99-777
; CURRENT APPLICATION NUMBER: US/09/155,077
; CURRENT FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Artificial/Unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(105)
; OTHER INFORMATION: Synthetic 105 mer peptide
US-09-155-077-1

Query Match 96.1%; Score 540; DB 15; Length 105;
Best Local Similarity 96.2%; Pred. No. 2.le-58;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 1 LYXKNRYALKSGSVNAPMPGVNTENNDDWVFMGYTQEEAKKNAMHKNQRIISGFSG 60
Db 2 LYXKNRYALKSGSVNAPMPGVNTENNDDWVFMGYTQEEAKKNAMHKNQRIISGFSG 61
Qy 61 PFGEENGKNGALNLFNGKSAQNRFLLTGCTNLNGKISVTQG 104
Db 62 PFGEENGKNGALNLFNGKSAQNRFLLTGCTNLNGKISVTQG 105

RESULT 5

US-08-578-788-4
; Sequence 4, Application US/08578788
; GENERAL INFORMATION:
; APPLICANT: KILIAN, Mogens
; APPLICANT: POULSEN, Knud
; TITLE OF INVENTION: IMMUNOGLOBULIN A1 PROTEASES (IGAL
; TITLE OF INVENTION: PROTEASES), METHOD OF INDUCTION OF BROAD-SPECTRUM
; TITLE OF INVENTION: INHIBITORY ANTIBODIES USING GENE-TECHNOLOGICALLY PRODUCED
; TITLE OF INVENTION: BACTERIAL IGAL PROTEASES APPLIED ALONE OR CONJUGATED TO BA
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN
; STREET: 400 Seventh Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578,788
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,638
; FILING DATE: 22-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: PRICE, D. DOUGLAS
; REGISTRATION NUMBER: 24,514

; REFERENCE/DOCKET NUMBER: 162/P54307A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-6666
; TELEFAX: (202) 393-5350
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1560 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-578-788-4

Query Match 96.1%; Score 540; DB 9; Length 1560;
Best Local Similarity 96.2%; Pred. No. 7.8e-57;
Matches 100; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 LYKKNRYALKSGGSYNAPMPGVNTENNNDWVPMGYTQEEAKKNAMNHNKNNRISGFSG 60
Db 584 LYKKNRYALKSGGSYNAPMPGVNTENNNDWVPMGYTQEEAKKNAMNHNKNNRISGFSG 643
QY 61 FFEENGKGHNALNLFNGKSAQNRELLTGGTNLNGKISVTQG 104
Db 644 FFEENGKGHNALNLFNGKSAQNRELLTGGTNLNGKISVTQG 687

RESULT 6

US-09-142-970-5
; Sequence 5, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtman
; APPLICANT: Monique Moreau
; TITLE OF INVENTION: IGAL PROTEASE FRAGMENT AS CARRIER
; FILE REFERENCE: 7101/0E6160S0
; CURRENT APPLICATION NUMBER: US/09/142,970
; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-09-142-970-5

Query Match 91.5%; Score 514; DB 15; Length 104;
Best Local Similarity 89.4%; Pred. No. 3.2e-55;
Matches 93; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 1 LYKKNRYALKSGGSYNAPMPGVNTENNNDWVPMGYTQEEAKKNAMNHNKNNRISGFSG 60
Db 1 LYKKNRYALKSGGRLNAPMPGVNTENNNDWVPMGYTQEEAKKNAMNHNKNNRISGDFG 60
QY 61 FFEENGKGHNALNLFNGKSAQNRELLTGGTNLNGKISVTQG 104
Db 61 FFEENGKGHNALNLFNGKSAQNRELLTGGTNLNGKISVTQG 104

RESULT 7

US-09-142-970-4
; Sequence 4, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtman
; APPLICANT: Monique Moreau
; TITLE OF INVENTION: IGAL PROTEASE FRAGMENT AS CARRIER
; FILE REFERENCE: 7101/0E6160S0
; CURRENT APPLICATION NUMBER: US/09/142,970

; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-09-142-970-4

Query Match 84.5%; Score 475; DB 15; Length 104;
Best Local Similarity 88.7%; Pred. No. 2e-50;
Matches 86; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 1 LYKKNRYALKSGGSYNAPMPGVNTENNNDWVPMGYTQEEAKKNAMNHNKNNRISGFSG 60
Db 1 LYKKNRYALKSGGRLNAPMPGVNTENNNDWVPMGYTQEEAKKNAMNHNKNNRISGDFG 60
QY 61 FFEENGKGHNALNLFNGKSAQNRELLTGGTNLNG 97
Db 61 FFEENGKGHNALNLFNGKSAQNRELLTGGTNLNG 97

RESULT 8

US-08-578-788-2
; Sequence 2, Application US/08578788
; GENERAL INFORMATION:
; APPLICANT: KILIAN, Moegens
; APPLICANT: POULSEN, Knud
; TITLE OF INVENTION: IMMUNOGLOBULIN A1 PROTEASES (IGAL
; TITLE OF INVENTION: PROTEASES), METHOD OF INDUCTION OF BROAD-SPECTRUM
; TITLE OF INVENTION: INHIBITORY ANTIBODIES USING GENE-TECHNOLOGICALLY PRODUCED
; TITLE OF INVENTION: BACTERIAL IGAL PROTEASES APPLIED ALONE OR CONJUGATED TO BA
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN
; STREET: 400 Seventh Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578,788
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,638
; FILING DATE: 22-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: PRICE, D. DOUGLAS
; REGISTRATION NUMBER: 24,514
; REFERENCE/DOCKET NUMBER: 162/P54307A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-6666
; TELEFAX: (202) 393-5350
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1540 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-578-788-2

Query Match

58.6%; Score 329.5; DB 9; Length 1540;

Best Local Similarity 56.7%; Pred. No. 5.5e-31;
Matches 59; Conservative 20; Mismatches 24; Indels 1; Gaps 1;

QY 1 LYKKNRYALKSGSVNAPMPGVNTENNDFVGYTQEEAKKNAMHKNORISGFSG 60
DB 595 LNLNYYTYALKRGASTSELPKNSGESNENWLYMGKTSDEAKRVNMHNNRNMNGFNG 654
QY 61 PFGEENGKNGALNLFNGKSAQNRFLLTGTGTLNKGISVTQG 104
DB 655 YFGEDEGK-NNGNLNVTFKSGEQNRFLLTGTGTLNGLDITVEKG 697

RESULT 9

US-08-296-791-3
; Sequence 3, Application US/08296791
; GENERAL INFORMATION:
; APPLICANT: St. Geme III, Joseph W.
; APPLICANT: Falkow, Stanley
; TITLE OF INVENTION: Haemophilus Adherence and Penetration
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1541 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
US-08-296-791-3

Query Match 58.6%; Score 329.5; DB 6; Length 1541;

Best Local Similarity 56.7%; Pred. No. 5.5e-31;
Matches 59; Conservative 20; Mismatches 24; Indels 1; Gaps 1;

QY 1 LYKKNRYALKSGSVNAPMPGVNTENNDFVGYTQEEAKKNAMHKNORISGFSG 60
DB 595 LNLNYYTYALKRGASTSELPKNSGESNENWLYMGKTSDEAKRVNMHNNRNMNGFNG 654
QY 61 PFGEENGKNGALNLFNGKSAQNRFLLTGTGTLNKGISVTQG 104
DB 655 YFGEDEGK-NNGNLNVTFKSGEQNRFLLTGTGTLNGLDITVEKG 697

RESULT 10

US-08-296-791-5
; Sequence 5, Application US/08296791
; GENERAL INFORMATION:
; APPLICANT: St. Geme III, Joseph W.
; APPLICANT: Falkow, Stanley
; TITLE OF INVENTION: Haemophilus Adherence and Penetration

; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1702 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
US-08-296-791-5

Query Match 57.4%; Score 322.5; DB 6; Length 1702;
Best Local Similarity 55.8%; Pred. No. 4.5e-30;
Matches 58; Conservative 19; Mismatches 26; Indels 1; Gaps 1;

QY 1 LYKKNRYALKSGSVNAPMPGVNTENNDFVGYTQEEAKKNAMHKNORISGFSG 60
DB 601 LNLNYYTYALKRGASTSELPKNSGESNENWLYMGKTSDEAKRVNMHNNRNMNGFNG 660
QY 61 PFGEENGKNGALNLFNGKSAQNRFLLTGTGTLNKGISVTQG 104
DB 661 YFGEDEGK-NNGNLNVTFKSGEQNRFLLTGTGTLNGLDITVEKG 703

RESULT 11

US-08-296-791-4
; Sequence 4, Application US/08296791
; GENERAL INFORMATION:
; APPLICANT: St. Geme III, Joseph W.
; APPLICANT: Falkow, Stanley
; TITLE OF INVENTION: Haemophilus Adherence and Penetration
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:


```

QY 1 1 LYX- -NRYVALKSGSVNAPENGVNTENNDDWVFGYQTEAKNMMHKNNQRISGF 58
   ||:  |  ||  ||  |  :  ||  |  :  ||  |  :  ||  |  :  ||  |  :  ||  |  :  ||
Db 603 LYFNQDNRSYTYLLKGASTRSELPQSGESNENWLYMGRSTDAAKRVNMHINRMNGF 662

QY 59 SGFPGENGKG-HNGALNLFNFKSAQNRELLTGGTNLNGKISVTOG 104
   ||:|||||  |  ||  ||:  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 663 NGYFGEETRATQNGKLVNFTNGKSDQNRELLTGGTNLNGDLNVEKG 709

RESULT 13
US-08-296-791-2
; Sequence 2, Application US/08296791
; GENERAL INFORMATION:
; APPLICANT: St. Geme III, Joseph W.
; APPLICANT: Falkow, Stanley
; TITLE OF INVENTION: Haemophilus adherence and Penetration
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1394 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-296-791-2

Query Match 20.3%; Score 114; DB 6; Length 1394;
Best Local Similarity 38.3%; Pred. No. 0.00014;
Matches 23; Conservative 13; Mismatches 24; Indels 0; Gaps

QY 45 NAMNHNKNQRISGFFGFGENGKGHNGALNLFNFKSAQNRELLTGGTNLNGKISVTOG 104
   |  :  |  :  :  :  ||  ||  |  ||  ||  |  :  :  :  :  :  :  :  :  :  :  :  :
Db 551 NNINKLDYREIAYNGWFGETDKNKGRLNLIYKPTTDRITLLSGGTNLKGDITQTKG 610

RESULT 14
US-09-733-089-13711
; Sequence 13711, Application US/09733089
; GENERAL INFORMATION:
; APPLICANT: Dotson, Stanton B.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Lutfiyya, Linda L.
; APPLICANT: McIninch, James
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated
; TITLE OF INVENTION: Transcription In Plants
; FILE REFERENCE: 38-21(15300)D

```

Query Match	55.6%	Score 312.5;	DB 6;	Length 1848;	
Best Local Similarity	54.2%;	Pred. No. 8.6e-29;			
Matches 58;	Conservative 18;	Mismatches 28;	Indels 3;	Gaps 2;	
<hr/>					
					APPLICANT: MCIninch, James
					APPLICANT: Wu, Wei
					TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
					TITLE OF INVENTION: Transcription In Plants
					FILE REFERENCE: 38-21(15300)D

; CURRENT APPLICATION NUMBER: US/09/733,089
; CURRENT FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 09/474,435
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 24143
; SEQ ID NO 13711
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(450)
; OTHER INFORMATION: unsure at all xaa locations
US-09-733-089-13711

Query Match 15.8%; Score 89; DB 21; Length 450;
Best Local Similarity 25.08; Pred. No. 0.036;
Matches 30; Conservative 17; Mismatches 43; Indels 30; Gaps 4;

QY 8 YYALKSGGSVNAPENGVTENNNDWFMG---YTQEEAKKNAMNH-----KNNQRIS 56
|| : : : : : || : : : : : || : : : : :
Db 33 YYITVSFSIASQSVSKGKITQEKFWDLGNLNLVQKXXEKKTIIVHLIFMSLLTVAHQRS 92
|| : : : : : || : : : : : || : : : : :

QY 57 G-FSGF-----FGEENGKCHGNALNPNFGSAQNRFLLTGTTNLNG 97
|| : : : : : || : : : : : || : : : : :
Db 93 GEFIRFTESHGGDDVSGDGEHGGCHDDGSGAGSLGLNPNVMOQGEVTMOGGSLVSG 152

RESULT 15
US-60-161-932-2066
; Sequence 2066, Application US/60161932
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: The Drosophila Genome, Primary Nucleic
; TITLE OF INVENTION: Acid Sequences, Systems Containin the Nucleic Acid Sequences
; TITLE OF INVENTION: and Uses Thereof
; FILE REFERENCE: CLO00122
; CURRENT APPLICATION NUMBER: US/60/161,932
; CURRENT FILING DATE: 1999-10-28
; NUMBER OF SEQ ID NOS: 2626
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2066
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Drosophila
US-60-161-932-2066

Query Match 14.7%; Score 82.5; DB 23; Length 426;
Best Local Similarity 33.3%; Pred. No. 0.21;
Matches 20; Conservative 14; Mismatches 23; Indels 3; Gaps 2;

QY 20 PMPENGVTENNNDWFMGYTQEEAKKNAMHNKNNQRISG--PSGFFGEGCKGHNGALN 77
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : :
Db 183 PQDVRVFKDGFSGFIKPTKEAAHAHAIEHTHNSVHGNLVKCFWKGNG--GDNSSANNLN 241

Search completed: May 1, 2001, 15:02:50
Job time: 332 sec


```

; Sequence 28, Application US/09813408
; GENERAL INFORMATION:
; APPLICANT: Delagrave, Simon
; APPLICANT: Maris, Barry
; TITLE OF INVENTION: Methods For The Synthesis Of Polynucleotides And Combinatorial
; TITLE OF INVENTION: Of Polynucleotides
; FILE REFERENCE: HER0041
; CURRENT APPLICATION NUMBER: US/09/813,408
; CURRENT FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 28
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Bacillus halodurans
US-09-813-408-28

```

Best Local Similarity 21.2%; Pred. No. 14;
Matches 22; Conservative 17; Mismatches 29; Indels 36; Gaps 5;

QY 2 YKKNRY-----YALKSGSVNAPMPENGVTENN-----DWVFMGYTQEEAKN-A 46
Db 307 FFKEYQNLQIDIEGWSLASGENVKIALDTGDPNHPDKIKVIYNGYNAVEGNNFA 366

QY 47 MNHKNRISG-----FSGFFG-----EENGKG 69

Db 367 DKHGHGTHVAGAAATNNVTGTAGISWKSSEILPVKVLNDNGEG 410

RESULT 7

US-09-792-024-92

; Sequence 92, Application US/09792024

; GENERAL INFORMATION:

; APPLICANT: Roemer, Terry

; APPLICANT: Jiang, Bo

; APPLICANT: Boone, Charles

; APPLICANT: Bussey, Howard

; TITLE OF INVENTION: Gene Disruption Methodologies for Drug

; TITLE OF INVENTION: Targets Discovery

; FILE REFERENCE: 10182-004-999

; CURRENT APPLICATION NUMBER: US/09/792,024

; CURRENT FILING DATE: 2001-02-20

; NUMBER OF SEQ ID NOS: 490

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 92

; LENGTH: 409

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-792-024-92

Query Match

Best Local Similarity 10.1%; Score 56.5; DB 5; Length 409;

Matches 14; Conservative 7; Mismatches 21; Indels 3; Gaps 1;

QY 8 YYALKSGSVNAPMPENGVTENNNDVFMGYTQEEAKKNAMNHN 52

Db 101 YKLLKSGKPKGDLP---LQNLDLVMTGTFTIEEREKEKEENN 142

RESULT 8

US-09-308-453-2

; Sequence 2, Application US/09308453

; GENERAL INFORMATION:

; APPLICANT: Roche Diagnostics GMBH

; TITLE OF INVENTION: Recombinant collagenase type I from Clostridium histolyticum and

; TITLE OF INVENTION: for isolating cells and groups of cells

; FILE REFERENCE: BMID9924US

; CURRENT APPLICATION NUMBER: US/09/308,453

; CURRENT FILING DATE: 1999-09-09

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 1008

; TYPE: PRT

; ORGANISM: Clostridium histolyticum

US-09-308-453-2

Query Match

Best Local Similarity 10.1%; Score 56.5; DB 5; Length 1008;

Matches 18; Conservative 21; Mismatches 22; Indels 23; Gaps 5;

QY 8 YYALKSGSVNAPMPENGVTENNNDVFMGYTQEEAKKNAMNH-----KNNQRISGSGF 61

Db 814 YFDVKEDGDTIELPYSS---SSNFTLVYKEGDQ---NHTASGIDKNSKV----GT 862

QY 62 FGENGK-----GHNGALNLFN 79

Db 863 FKATKGRHYFYIKHDSASNISYS 886

RESULT 9

US-09-392-846-2

; Sequence 2, Application US/09392846

; GENERAL INFORMATION:

; APPLICANT: Pienkos, Philip T.

; APPLICANT: Kerridge, Alison

; APPLICANT: Holland, Herbert L.

; TITLE OF INVENTION: Method of Producing Chiral Sulfoxides

; FILE REFERENCE: 1405.1103-000

; CURRENT APPLICATION NUMBER: US/09/392,846

; CURRENT FILING DATE: 1999-09-09

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 417

; TYPE: PRT

; ORGANISM: Rhodococcus species

US-09-392-846-2

Query Match

Best Local Similarity 10.0%; Score 56; DB 5; Length 417;

Matches 11; Conservative 9; Mismatches 12; Indels 2; Gaps 1;

QY 12 KSGSVNAPMPEN--GVTEENNDFVFMGYTQEEAK 43

Db 183 QQGAITAAIPTSRAGVTPNDWAAIGMROTDSR 216

RESULT 10

US-09-804-472-2

; Sequence 2, Application US/09804472

; GENERAL INFORMATION:

; APPLICANT: SHAO, Wei et al.

; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,

; TITLE OF INVENTION: AND USES THEREOF

; FILE REFERENCE: CL001163

; CURRENT APPLICATION NUMBER: US/09/804,472

; CURRENT FILING DATE: 2001-03-13

; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2

; LENGTH: 791

; TYPE: PRT

; ORGANISM: Human

US-09-804-472-2

Query Match

Best Local Similarity 10.0%; Score 56; DB 5; Length 791;

Matches 14; Conservative 16; Mismatches 20; Indels 12; Gaps 2;

QY 3 YKNRYVALKSGSVNA-----PMPENGVTENNDFVFMGYTQEEAKKNAMNHN 52

Db 24 HKRGTHYWTNGGINSSTHLLDLLDEPIP--GVGYDDFTIDWVREKCKDRHRHRRIN 81

QY 53 QR 54

Db 82 SK 83

RESULT 11

US-60-248-823-105

; Sequence 105, Application US/60248823

; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen

; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL000949

; CURRENT APPLICATION NUMBER: US/60/248,823
; CURRENT FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 1455
; TYPE: PRT
; ORGANISM: HUMAN
US-60-248-823-105

Query Match 9.8%; Score 55; DB 6; Length 1455;
Best Local Similarity 19.8%; Pred. No. 80;
Matches 16; Conservative 17; Mismatches 42; Indels 6; Gaps 1;

QY 17 VNAPMPGVCTENNNDWFMGYTQEEAKKNAMHKNQRIISGFSFGFGKNGHNGALNL 76
DB 771 LSEPVQTNHESKEWYIASLTRAQAEHMLM-----RVPRDGAFLVRRKNEPNSYISF 824
QY 77 NFNKSAQNRFLLTGGTNLNG 97
DB 825 RAEGKIKHCRVQEGQTVMLG 845

RESULT 12
US-09-739-449-12040
; Sequence 12040, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 12040
; LENGTH: 124
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-12040

Query Match 9.7%; Score 54.5; DB 5; Length 124;
Best Local Similarity 32.4%; Pred. No. 4.1;
Matches 12; Conservative 12; Mismatches 8; Indels 5; Gaps 2;

QY 66 NKGHNGALNLFNFKSAQNRFLLTGGTNLNGKISVT 102
DB 81 NGRGYHAGILKSMGKTAQ----ILGG-NQSGRVQVS 112

RESULT 13
US-60-248-505-876
; Sequence 876, Application US/60248505
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: c1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 876
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Human
US-60-248-505-876

Query Match 9.7%; Score 54.5; DB 6; Length 445;
Best Local Similarity 29.5%; Pred. No. 21;
Matches 18; Conservative 8; Mismatches 24; Indels 11; Gaps 2;
QY 20 PMPENGV-----TENNNDWFMGYTQEEAKKNAMHKNQRIISGFSFGFGKNGH 70
DB 25 PLPVGVDPDTCQSTPTRSDPVLLMHT--NFKTNAENAGKNKTIROKNGNNGHPNANGH 82
QY 71 N 71
DB 83 N 83

RESULT 14
US-09-193-562D-11
; Sequence 11, Application US/09193562D
; GENERAL INFORMATION:
; APPLICANT: Pauli, Benedicht U.
; TITLE OF INVENTION: Nucleotide Sequences Encoding Mammalian Calcium
; TITLE OF INVENTION: Activated Chloride Channel-Adhesion Molecules
; FILE REFERENCE: 18617.0052
; CURRENT APPLICATION NUMBER: US/09/193,562D
; CURRENT FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US/60/065,922
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 11
; LENGTH: 795
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Variant of Lu-ECAM-1 from bovine endothelial cells
US-09-193-562D-11

Query Match 9.7%; Score 54.5; DB 5; Length 795;
Best Local Similarity 25.8%; Pred. No. 43;
Matches 25; Conservative 10; Mismatches 31; Indels 31; Gaps 5;

QY 12 KSGSVNAPMPENGV---TENNDWFMGYTQEEAKKNAMHKNQRIISGFSFGFGENG- 67
DB 641 EDGHQVTLLEWDNGAGRTVNDGIYSY-----FTDYYG--NGR 678
QY 68 ---KGHNGALN--LNFNGKSAQNRFLLTGGTNLNGKI 99
DB 679 YSLKVHAQARNNTARLNLRQPNQKVLVYVPGYVNGKI 715

RESULT 15
US-09-193-562D-12
; Sequence 12, Application US/09193562D
; GENERAL INFORMATION:
; APPLICANT: Pauli, Benedicht U.
; TITLE OF INVENTION: Nucleotide Sequences Encoding Mammalian Calcium
; TITLE OF INVENTION: Activated Chloride Channel-Adhesion Molecules
; FILE REFERENCE: 18617.0052
; CURRENT APPLICATION NUMBER: US/09/193,562D
; CURRENT FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: US/60/065,922
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 12
; LENGTH: 821
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Variant of Lu-ECAM-1 from bovine endothelial cells
US-09-193-562D-12

Query Match 9.7%; Score 54.5; DB 5; Length 821;
Best Local Similarity 25.8%; Pred. No. 44;
Matches 25; Conservative 10; Mismatches 31; Indels 31; Gaps 5;

QY 12 KSGSVNAPMPENG- ---TENNDWVFMGYTQEEAKKNAMNHNQRIISGFGFGEENG- 67
Db 641 EDGHQVTLLELDNGAGRDYKNDGIYSRY-----FTDIYG--NGR 678
QY 68 ---KGHNALN--LNFNGKSAQNRELLTGGTNLNGKI 99
Db 679 YSLKVHAQARNNTARLNLROPONKVLVPGYVENGKI 715

Search completed: May 1, 2001, 15:03:03
Job time: 319 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 15:02:50 ; Search time 149.44 Seconds
(without alignments)
111.944 Million cell updates/sec

Title: US-09-142-970-3
Perfect score: 563
Sequence: 1 LYKKNRYALKSGSVNAP.....NRFLTGTGTLNGKISVTQG 104

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1009251 seqs, 16085430 residues
Total number of hits satisfying chosen parameters: 1009251

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main.*
1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US080_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	563	100.0	104	15	US-09-142-970-3
2	546	97.0	104	15	US-09-142-970-2
3	537	95.4	104	15	US-09-142-970-1
4	537	95.4	105	15	US-09-155-077-1
5	537	95.4	1560	9	US-08-578-788-4
6	501	89.0	104	15	US-09-142-970-5
7	462	82.1	104	15	US-09-142-970-4
8	326.5	58.0	1545	6	US-08-296-791-4
9	320.5	56.9	1540	9	US-08-578-788-2
10	320.5	56.9	1541	6	US-08-296-791-3

11	313.5	55.7	1702	6	US-08-296-791-5	Sequence 5, Appli
12	303.5	53.9	1848	6	US-08-296-791-6	Sequence 6, Appli
13	114	20.2	1394	6	US-08-296-791-2	Sequence 2, Appli
14	84	14.9	1201	23	US-60-167-217-7969	Sequence 7969, Ap
15	84	14.9	1201	23	US-60-173-464-6367	Sequence 6367, Ap
16	84	14.9	1225	23	US-60-191-637-7869	Sequence 7869, Ap
17	84	14.9	1225	23	US-60-191-681-6135	Sequence 6135, Ap
18	82	14.6	450	21	US-09-733-089-13711	Sequence 13711, A
19	79	14.0	402	23	US-60-191-637-28213	Sequence 28213, A
20	79	14.0	402	23	US-60-191-681-22822	Sequence 22822, A
21	79	14.0	658	13	US-08-993-002A-9518	Sequence 9518, Ap
22	79	14.0	996	11	US-08-761-066-544	Sequence 544, App
23	79	14.0	996	13	US-08-993-002A-9519	Sequence 9519, Ap
24	77	13.7	229	16	US-09-248-796-24831	Sequence 24831, A
25	76.5	13.6	280	1	PCT-US00-09066-123	Sequence 123, App
26	76.5	13.6	954	19	US-09-570-856-41	Sequence 41, Appl
27	76	13.5	419	16	US-09-270-767-41767	Sequence 41767, A
28	76	13.5	586	23	US-60-167-217-15320	Sequence 15320, A
29	76	13.5	586	23	US-60-171-625-324	Sequence 324, App
30	76	13.5	586	23	US-60-173-464-12542	Sequence 12542, A
31	76	13.5	586	23	US-60-191-637-15302	Sequence 15302, A
32	76	13.5	586	23	US-60-191-637-28261	Sequence 28261, A
33	76	13.5	586	23	US-60-191-681-12116	Sequence 12116, A
34	76	13.5	586	23	US-60-191-681-22870	Sequence 22870, A
35	76	13.5	586	23	US-60-219-005-174	Sequence 174, App
36	76	13.5	586	23	US-60-219-005-297	Sequence 297, App
37	74.5	13.2	426	23	US-60-161-932-2066	Sequence 2066, Ap
38	74.5	13.2	464	23	US-60-167-217-11345	Sequence 11345, A
39	74.5	13.2	464	23	US-60-167-217-11363	Sequence 11363, A
40	74.5	13.2	464	23	US-60-173-464-9184	Sequence 9184, Ap
41	74.5	13.2	464	23	US-60-173-464-9196	Sequence 9196, Ap
42	74.5	13.2	464	23	US-60-191-637-11351	Sequence 11351, A
43	74.5	13.2	464	23	US-60-191-637-11369	Sequence 11369, A
44	74.5	13.2	464	23	US-60-191-681-8894	Sequence 8894, Ap
45	74.5	13.2	464	23	US-60-191-681-8906	Sequence 8906, Ap

ALIGNMENTS

RESULT 1
US-09-142-970-3
; Sequence 3, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtman
; APPLICANT: Monique Moreau
; TITLE OF INVENTION: IGAI PROTEASE FRAGMENT AS CARRIER
; FILE REFERENCE: 7101/08616USO
; CURRENT FILING DATE: 1999-04-02
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-142-970-3

Query Match 100.0%; Score 563; DB 15; Length 104;
Best Local Similarity 100.0%; Pred. No. 8.5e-61;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LYKKNRYALKSGSVNAPMPENGOTENDWFMGYKOEAKNMMHKNORISGFSG 60
|||||
Db 1 LYKKNRYALKSGSVNAPMPENGOTENDWFMGYKOEAKNMMHKNORISGFSG 60
QY 61 FFGEENGKGNALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 104

Db 61 FFGENGKGHNGALNLFNGKSAQNRELLTGGTNLNGKISVTQ 104
|||||

RESULT 2

US-09-142-970-2
; Sequence 2, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtmann
; APPLICANT: Monique Moreau
; TITLE OF INVENTION: IGAL PROTEASE FRAGMENT AS CARRIER
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE: 7101/0E616U50
; CURRENT APPLICATION NUMBER: US/09/142.970
; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-142-970-2

Query Match 97.08; Score 546; DB 15; Length 104;
Best Local Similarity 97.18; Pred. No. 1e-58; Indels 0; Gaps 0;
Matches 101; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LYKNRYALKSGSVNAPMPENGQTNNDWFMGYKQEEAKNMMHKNQRIISGFSG 60
|||||

Db 1 LYKNRYALKSGSVNAPMPENGQTNNDWFMGYTQEEAKNMMHKNQRIISGFSG 60
|||||

RESULT 3

US-09-142-970-1
; Sequence 1, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtmann
; APPLICANT: Monique Moreau
; TITLE OF INVENTION: IGAL PROTEASE FRAGMENT AS CARRIER
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE: 7101/0E616U50
; CURRENT APPLICATION NUMBER: US/09/142.970
; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-142-970-1

Query Match 95.48; Score 537; DB 15; Length 104;
Best Local Similarity 95.28; Pred. No. 1.3e-57; Indels 0; Gaps 0;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LYKNRYALKSGSVNAPMPENGQTNNDWFMGYKQEEAKNMMHKNQRIISGFSG 60
|||||

Db 1 LYKNRYALKSGSVNAPMPENGQTNNDWILMGSTQEEAKNMMHKNQRIISGFSG 60
|||||

Qy 61 FFGENGKGHNGALNLFNGKSAQNRELLTGGTNLNGKISVTQ 104
|||||

Db 61 FFGENGKGHNGALNLFNGKSAQNRELLTGGTNLNGKISVTQ 104
|||||

RESULT 4

US-09-155-077-1
; Sequence 1, Application US/09155077
; GENERAL INFORMATION:
; APPLICANT: Moreau, Monique
; APPLICANT: Mistretta, Noelle
; TITLE OF INVENTION: Polysaccharide-Peptide-Conjugates
; FILE REFERENCE: MBHB99-777
; CURRENT APPLICATION NUMBER: US/09/155.077
; CURRENT FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Artificial/Unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(105)
; OTHER INFORMATION: Synthetic 105 mer peptide
US-09-155-077-1

Query Match 95.48; Score 537; DB 15; Length 105;
Best Local Similarity 95.28; Pred. No. 1.3e-57; Indels 0; Gaps 0;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LYKNRYALKSGSVNAPMPENGQTNNDWFMGYKQEEAKNMMHKNQRIISGFSG 60
|||||

Db 2 LYKNRYALKSGSVNAPMPENGQTNNDWILMGSTQEEAKNMMHKNQRIISGFSG 61
|||||

RESULT 5

US-08-578-788-4
; Sequence 4, Application US/08578788
; GENERAL INFORMATION:
; APPLICANT: KILIAN, Mogens
; APPLICANT: POULSEN, Knud
; TITLE OF INVENTION: IMMUNOGLOBULIN A1 PROTEASES (IGAL
; TITLE OF INVENTION: PROTEASES), METHOD OF INDUCTION OF BROAD-SPECTRUM
; TITLE OF INVENTION: INHIBITORY ANTIBODIES USING GENE-TECHNOLOGICALLY PRODUCED
; TITLE OF INVENTION: BACTERIAL IGAL PROTEASES APPLIED ALONE OR CONJUGATED TO BA
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN
; STREET: 400 Seventh Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578,788
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,638
; FILING DATE: 22-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: PRICE, D. DOUGLAS
; REGISTRATION NUMBER: 24,514

Query Match 95.48; Score 537; DB 15; Length 104;
Best Local Similarity 95.28; Pred. No. 1.3e-57; Indels 0; Gaps 0;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LYKNRYALKSGSVNAPMPENGQTNNDWFMGYKQEEAKNMMHKNQRIISGFSG 60
|||||

Db 2 LYKNRYALKSGSVNAPMPENGQTNNDWILMGSTQEEAKNMMHKNQRIISGFSG 61
|||||

REFERENCE/DOCKET NUMBER: 162/P54307A
TELEPHONE: (202) 638-6666
TELEFAX: (202) 393-5350
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1560 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-578-788-4

Query Match 95.4%; Score 537; DB 9; Length 1560;
Best Local Similarity 95.2%; Pred. No. 4.9e-56;
Matches 99; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 LYKKNRYALKSGGSYNAPMPENGQTNNDWFMGYKQEEAKNAMHNKNNRISGFSG 60
Db 584 LYKKNRYALKSGGSYNAPMPENGQTNNDWFMGYKQEEAKNAMHNKNNRISGFSG 643
QY 61 FFEENGKGHGALNLFNFKSAQNRFLLTGGTNLKGKISVTG 104
Db 644 FFEENGKGHGALNLFNFKSAQNRFLLTGGTNLKGKISVTG 687

RESULT 6

US-09-142-970-5
Sequence 5, Application US/09142970
GENERAL INFORMATION:
APPLICANT: Mark Achtman
TITLE OF INVENTION: IGAL PROTEASE FRAGMENT AS CARRIER
FILE REFERENCE: 7101/056160S
CURRENT APPLICATION NUMBER: US/09/142,970
EARLIER FILING DATE: 1999-04-02
EARLIER FILING DATE: 1997-01-21
EARLIER FILING DATE: 1997-01-21
EARLIER FILING DATE: 1998-01-20
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5
LENGTH: 104
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-09-142-970-5

Query Match 89.0%; Score 501; DB 15; Length 104;
Best Local Similarity 87.5%; Pred. No. 3.1e-53;
Matches 91; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
QY 1 LYKKNRYALKSGGSYNAPMPENGQTNNDWFMGYKQEEAKNAMHNKNNRISGFSG 60
Db 1 LYKKNRYALKSGGRLNAPMPENGVAENNDWFMGYTQEEAKNAMHNKNNRIGDFG 60
QY 61 FFEENGKGHGALNLFNFKSAQNRFLLTGGTNLKGKISVTG 104
Db 61 FFEENGKGHGALNLFNFKSAQNRFLLTGGTNLKGKISVTG 104

RESULT 7

US-09-142-970-4
Sequence 4, Application US/09142970
GENERAL INFORMATION:
APPLICANT: Mark Achtman
TITLE OF INVENTION: IGAL PROTEASE FRAGMENT AS CARRIER
FILE REFERENCE: 7101/056160S
CURRENT APPLICATION NUMBER: US/09/142,970

CURRENT FILING DATE: 1999-04-02
EARLIER APPLICATION NUMBER: EP 97100883.4
EARLIER FILING DATE: 1997-01-21
EARLIER APPLICATION NUMBER: PCT/EP98/00294
EARLIER FILING DATE: 1998-01-20
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 104
TYPE: PRT
ORGANISM: Neisseria gonorrhoeae
US-09-142-970-4

Query Match 82.1%; Score 462; DB 15; Length 104;
Best Local Similarity 86.6%; Pred. No. 1.8e-48;
Matches 84; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
QY 1 LYKKNRYALKSGGSYNAPMPENGQTNNDWFMGYKQEEAKNAMHNKNNRISGFSG 60
Db 1 LYKKNRYALKSGGRLNAPMPENGVAENNDWFMGYTQEEAKNAMHNKNNRIGDFG 60
QY 61 FFEENGKGHGALNLFNFKSAQNRFLLTGGTNLNG 97
Db 61 FFEENGKGHGALNLFNFKSAQNRFLLTGGANLNG 97

RESULT 8

US-08-296-791-4
Sequence 4, Application US/08296791
GENERAL INFORMATION:
APPLICANT: St. Geme III, Joseph W.
APPLICANT: Falkow, Stanley
TITLE OF INVENTION: Haemophilus Adherence and Penetration
TITLE OF INVENTION: Protein
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert
CITY: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/296,791
FILING DATE: 25-AUG-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Treccartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-59941/RPT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1545 amino acids
TYPE: amino acid
TOPOLOGY: unknown
US-08-296-791-4

Query Match 58.0%; Score 326.5; DB 6; Length 1545;
Best Local Similarity 57.5%; Pred. No. 2.3e-30;
Matches 61; Conservative 19; Mismatches 23; Indels 3; Gaps 2;
QY 1 LXY--KNRYALKSGGSYNAPMPENGQTNNDWFMGYKQEEAKNAMHNKNNRISGF 58

```

; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1541 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; US-08-296-791-3

Query Match 56.9%; Score 320.5; DB 6; Length 1541;
Best Local Similarity 54.8%; Pred. No. 1.2e-29;
Matches 57; Conservative 21; Mismatches 25; Indels 1;

QY 1 LYXKNRYAYALKSGSVNAPENGQOTENNDDWVFGYKQEBAQKNAHMKNNRISG
| :|| ||||| :| :| :| | :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 595 LNLNNTYYALRKGASTRSELPKKSAGESNEWNLYMGKTSDAKRVNMHNINRMNG
||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
QY 61 FFGEENGKHGNLALNFNGSAQNRRFLTGTGTLNIGKISVTQG 104
||||| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :|| :||
Db 655 YFGECECK-NNGNLNVTFKGKSEQNRFLTGTGTLNGLDITVEKG 697

RESULT 11
US-08-296-791-5
; Sequence 5, Application US/08296791
; GENERAL INFORMATION:
; APPLICANT: St. Geme III, Joseph W.
; APPLICANT: Falkow, Stanley
; TITLE OF INVENTION: Haemophilus Adherence and Penetration
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```

[illegible]

Search completed: May 1, 2001, 15:02:51
Job time: 333 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 15:03:03 ; Search time 11.48 Seconds
(without alignments)
37.203 Million cell updates/sec

Title: US-09-142-970-3
Perfect score: 563
Sequence: 1 LYKNRYRYALKSGSVNAP.....NRFLTGTNLNGKISVTQG 104

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 25343 seqs, 4106610 residues

Total number of hits satisfying chosen parameters: 25343

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New.*

- 1: /cgn2.6/ptodata/1/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2.6/ptodata/1/paa/US06_NEW_COMB.pep.*
- 3: /cgn2.6/ptodata/1/paa/US07_NEW_COMB.pep.*
- 4: /cgn2.6/ptodata/1/paa/US08_NEW_COMB.pep.*
- 5: /cgn2.6/ptodata/1/paa/US09_NEW_COMB.pep.*
- 6: /cgn2.6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66.5	11.8	1008	US-09-308-453-2	Sequence 2, Appli
2	59	10.5	495	US-09-819-386-5	Sequence 5, Appli
3	59	10.5	495	US-09-819-386-2	Sequence 2, Appli
4	58.5	10.4	421	US-09-420-785A-4	Sequence 4, Appli
5	58	10.3	342	US-09-509-031-6	Sequence 6, Appli
6	58	10.3	495	US-09-509-031-4	Sequence 4, Appli
7	57.5	10.2	302	US-09-308-823A-216	Sequence 216, App
8	56.5	10.0	185	US-09-739-449-10404	Sequence 10404, A
9	55.5	9.9	1205	US-09-806-080-4	Sequence 4, Appli
10	55	9.8	444	PCT-US01-11988-807	Sequence 807, App
11	55	9.8	1272	US-09-739-449-10101	Sequence 10101, A
12	54.5	9.7	124	US-09-739-449-12040	Sequence 12040, A
13	54.5	9.7	397	US-09-739-449-13127	Sequence 13127, A
14	54.5	9.7	427	US-09-739-449-11055	Sequence 11055, A
15	54.5	9.7	795	US-09-193-562D-11	Sequence 11, Appli
16	54.5	9.7	821	US-09-193-562D-12	Sequence 12, Appli
17	54.5	9.7	905	US-09-193-562D-2	Sequence 2, Appli
18	54	9.6	704	US-09-646-351-21	Sequence 21, Appli
19	54	9.6	1833	US-08-945-567C-4	Sequence 4, Appli
20	54	9.6	1992	US-08-945-567C-3	Sequence 3, Appli
21	53.5	9.5	409	US-09-792-024-92	Sequence 92, Appli
22	53.5	9.5	551	US-09-818-143-21	Sequence 21, Appli
23	53.5	9.5	1976	US-09-739-449-9987	Sequence 9987, Ap
24	53	9.4	415	US-09-802-633-2	Sequence 2, Appli
25	52	9.2	287	PCT-US01-11797-30	Sequence 30, Appli
26	52	9.2	417	US-09-392-846-2	Sequence 2, Appli
27	52	9.2	791	US-09-804-472-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-308-453-2
; Sequence 2, Application US/09308453
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GmbH
; TITLE OF INVENTION: Recombinant collagenase type I from Clostridium histolyticum a
; FILE REFERENCE: BMID924US
; CURRENT APPLICATION NUMBER: US/09/308,453
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1008
; TYPE: PRT
; ORGANISM: Clostridium histolyticum
US-09-308-453-2

Query Match 11.8%; Score 66.5; DB 5; Length 1008;
Best Local Similarity 23.8%; Pred. No. 2;
Matches 20; Conservative 20; Mismatches 21; Indels 23; Gaps 5;

QY 8 YALKSGGSYNADPMPNGQTNNDWTFMGYKQEEAKNMNH-----KNNQRISGSGF 61
DB 814 YFDVKEDGDVTIELPYSG---SSNFTWLVYKGGDQ----NHIASGIDKNNSKV----GT 862

QY 62 FGEENGK-----GHNGALNLNFN 79
DB 863 FRATKGRHVYFYKHDSASNISYS 886

RESULT 2
US-09-819-386-5
; Sequence 5, Application US/09819386
; GENERAL INFORMATION:
; APPLICANT: THORPE, PHILIP E.
; APPLICANT: RAN, SOPHIA
; TITLE OF INVENTION: CANCER TREATMENT METHODS USING THERAPEUTIC CONJUGATES
; FILE REFERENCE: 4001 002300
; CURRENT APPLICATION NUMBER: US/09/819,386
; CURRENT FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US/09/351,457
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 495
; TYPE: PRT

Sequence 13178, A
Sequence 223, App
Sequence 8767, Ap
Sequence 9158, Ap
Sequence 13, Appli
Sequence 28, Appli
Sequence 5254, Ap
Sequence 1051, Ap
Sequence 757, App
Sequence 8924, Ap
Sequence 10317, A
Sequence 120, App
Sequence 1, Appli
Sequence 10143, A
Sequence 10060, A
Sequence 11506, A
Sequence 10732, A
Sequence 9878, Ap

; ORGANISM: Homo sapiens
US-09-819-386-5

Query Match 10.5%; Score 59; DB 5; Length 495;
Best Local Similarity 31.3%; Pred. No. 6.3;
Matches 21; Conservative 10; Mismatches 16; Indels 20; Gaps 4;
Qy 48 NHKNNQR--ISGFGFGEENGKGNHGNALNLFNGKSAQN-----RFLITGG----- 92
Db 393 NEKQNYRLYLGHTGTAGKQSSLLHGA---DFSTKDADNDNCCKCALMLTGGWFDAC 449
Qy 93 --TNLNG 97
Db 450 GPSNLNG 456

RESULT 3
US-09-819-386-2
; Sequence 2, Application US/09819386
; GENERAL INFORMATION:
; APPLICANT: THORPE, PHILIP E.
; APPLICANT: RAN, SOPHIA
; TITLE OF INVENTION: CANCER TREATMENT METHODS USING THERAPEUTIC CONJUGATES
; FILE REFERENCE: 4001.002300
; CURRENT APPLICATION NUMBER: US/09/819,386
; CURRENT FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US/09/351,457
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-819-386-2

Query Match 10.5%; Score 59; DB 5; Length 498;
Best Local Similarity 31.3%; Pred. No. 6.4;
Matches 21; Conservative 10; Mismatches 16; Indels 20; Gaps 4;
Qy 48 NHKNNQR--ISGFGFGEENGKGNHGNALNLFNGKSAQN-----RFLITGG----- 92
Db 396 NEKQNYRLYLGHTGTAGKQSSLLHGA---DFSTKDADNDNCCKCALMLTGGWFDAC 452
Qy 93 --TNLNG 97
Db 453 GPSNLNG 459

RESULT 4
US-09-420-785A-4
; Sequence 4, Application US/09420785A
; GENERAL INFORMATION:
; APPLICANT: MORTESEN, UFFE
; APPLICANT: OLESEN, KJELD
; APPLICANT: STENNICKE, HENNING
; APPLICANT: SORESEN, STEEN B.
; APPLICANT: BREDDAM, KLAUS
; TITLE OF INVENTION: MODIFIED CARBOXYPEPTIDASE
; FILE REFERENCE: 089187/0109
; CURRENT APPLICATION NUMBER: US/09/420,785A
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; TYPE: PRT
; LENGTH: 421
; ORGANISM: Saccharomyces cerevisiae
US-09-420-785A-4

Query Match 10.4%; Score 58.5; DB 5; Length 421;
Best Local Similarity 33.3%; Pred. No. 5.9;
Matches 12; Conservative 9; Mismatches 14; Indels 1; Gaps 1;
Qy 3 YKNRYRYALKSGG-SVNAPMPENGOTENNNDVWFMY 37
Db 384 YKHFTYLRVFNHGHVFPDVPENALSMVNEWHGGF 419

RESULT 5
US-09-509-031-6
; Sequence 6, Application US/09509031
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBOIDES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/09/509,031
; CURRENT FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 342
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: TLHL protein
US-09-509-031-6

Query Match 10.3%; Score 58; DB 5; Length 342;
Best Local Similarity 26.5%; Pred. No. 5.2;
Matches 26; Conservative 12; Mismatches 32; Indels 28; Gaps 5;
Qy 13 SGGSVNAPMPENGOTENNNDVWFMYKOEAA--QKNAMNHNKRNORISGSGFF----- 62
Db 151 SGGSGSGSGSGSGSEN--LYFGGSAEEVITKANLIFANGSTOTAFKGTFFKATSEAY 208
Qy 63 -----GEENG-----KGHNALNLFNGKSAQNR 86
Db 209 AYADTLKDNCEYTVDVADKGY--TLNIKFKAGKEATNR 244

RESULT 6
US-09-509-031-4
; Sequence 4, Application US/09509031
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBOIDES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/09/509,031
; CURRENT FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: CATAB-TEV
US-09-509-031-4

Query Match 10.3%; Score 58; DB 5; Length 495;
Best Local Similarity 26.5%; Pred. No. 8.3;


```
Matches 26; Conservative 12; Mismatches 32; Indels 28; Gaps 5;

QY 13 SGGSVNAPMPENGOTENNDFVFMGYKOEAA--QKNAMNHKNRISGSGFF-----62
Db 151 SGGGSGGGSGGSGSEN--LYFQGSAREEYTIKANLIFANGSTQAEFKGTFEKATSEAY 208
QY 63 -----GEENG-----KGHNALNLFNGKSAQNR 86
Db 209 AYADTLAKDNGEYTVADVADKGY--TLNKFAGKEATNR 244

RESULT 7
US-09-308-823A-216
; Sequence 216, Application US/09308823A
; GENERAL INFORMATION:
; APPLICANT: Warren, Richard
; TITLE OF INVENTION: Novel Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; FILE REFERENCE: P50596
; CURRENT APPLICATION NUMBER: US/09/308, 823A
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ ID NOS: 661
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 216
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-308-823A-216

Query Match 10.2%; Score 57.5; DB 5; Length 302;
Best Local Similarity 28.9%; Pred. No. 5.1;
Matches 24; Conservative 11; Mismatches 35; Indels 13; Gaps 5;

QY 12 KSGSVNAPMPENGOTENNDFVFMGY-KQEEAQKNAMNHKNRISG--TSG-----PFGE 64
Db 226 KGGGSGS-----GGYNEGNQDYGCGFGGQSGQNNSSNTKQSDNPFSGGQYFGG 281
QY 65 ENKGHNGALNLFNGKSAQNR 87
Db 282 QNDNPFGG--NQNNQRNQGNSF 302

RESULT 8
US-09-739-449-10404
; Sequence 10404, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739, 449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514, 000
; PRIOR FILING DATE: 2000-02-23
; NUMBER OF SEQ ID NOS: 13351
; SEQ ID NO 10404
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-10404

Query Match 10.0%; Score 56.5; DB 5; Length 185;
Best Local Similarity 20.6%; Pred. No. 3.7;
Matches 13; Conservative 13; Mismatches 32; Indels 5; Gaps 1;

QY 14 GGSVNAPMPENGOTENNDFVFMGYKOEAAQKNAMNHKNRISG-----SGFFGEENGK 68
Db 95 QGNLRPLPELNEERRKSLVKVAHEYSEKAKVAIRHVRDRCDGLKAEKDGIDQDES 154
QY 69 GHN 71
```

```
Db 155 GQS 157

RESULT 9
US-09-806-080-4
; Sequence 4, Application US/09806080
; GENERAL INFORMATION:
; APPLICANT: Jomaa, Hassan
; TITLE OF INVENTION: Gene des 1-Desoxy-D-xylulose-Biosynthesewegs
; FILE REFERENCE: 15696
; CURRENT APPLICATION NUMBER: US/09/806, 080
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: DE199233567.8
; PRIOR FILING DATE: 1999-05-22
; PRIOR APPLICATION NUMBER: DE19843279.8
; PRIOR FILING DATE: 1998-09-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1205
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
US-09-806-080-4

Query Match 9.9%; Score 55.5; DB 5; Length 1205;
Best Local Similarity 26.1%; Pred. No. 49;
Matches 31; Conservative 10; Mismatches 51; Indels 27; Gaps 6;

QY 2 YY-----KNRYALKSGGSVNAPMPENGOTENNDFVFMGYKOEAAQKNAMNHKNRIS 55
Db 542 YYEAEWQVKNKEY-----GNGDIEISDNANVTNNERIFO--KGIHNDNNNNNNNNY 594
QY 56 SGFSGFGGEGNGHNGALNLFNGKSAQNRFL-----LTGG-----TNLNGKISV 101
Db 595 INPDSVVGRENTNPN-VNRDNHNVKDHVIAIIGDGLTGTGMALEALNYISFLSKILI 652

RESULT 10
PCT-US01-11988-807
; Sequence 807, Application PC/TUS0111988
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: PCT/US01/11988
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 807
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (92)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (95)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (97)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (98)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
```

Query Match	9.7%	Score 54.5;	DB 5;	Length 124;
Best Local Similarity	32.4%	Pred. No. 3.8;		

```

Query Match          9.7%; Score 54.5; DB 5; Length 427;
Best Local Similarity 33.3%; Pred. No.18;
Matches 19; Conservative 8; Mismatches 19; Indels 11; Gaps 4;

Qy 16 SVNAPMPENGOTE--NNDWFMVGYKQEAQKAMN--HKNNORISGFSGFFGEECKG 69
      | : | | : | | | : | | : | | | : | | |
Db 321 SLKRLFTQLGDTFECEV---YQGIGMDNAVPRFHKFAENVIGFSGY----NGRG 369

RESULT 15
US-09-193-562D-11
; Sequence 11, Application US/09193562D
; GENERAL INFORMATION:

```

APPLICANT: Pauli, Benedicht U.
TITLE OF INVENTION: Nucleotide Sequences Encoding Mammalian Calcium
FILE OF INVENTION: Activated Chloride Channel-Adhesion Molecules
FILE REFERENCE: 18617.0052
CURRENT APPLICATION NUMBER: US/09/193,562D
CURRENT FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: US/60/065,922
PRIOR FILING DATE: 1997-11-17
NUMBER OF SEQ ID NOS: 47
SEQ ID NO 11
LENGTH: 795
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Variant of Lu-ECAM-1 from bovine endothelial cells
US-09-193-562D-11

Query Match 9.7%; Score 54.5; DB 5; Length 795;
Best Local Similarity 25.8%; Pred. No. 39;
Matches 25; Conservative 10; Mismatches 31; Indels 31; Gaps 5;
QY 12 KSGGSVNAPENG--QTENNDFVFMGYKQEEAQKNAMNHNKNNQRISGFSGFGEENG- 67
Db 641 EDGHQVTELDWDNGAGRDTVKNDGIYSRY-----FTDIYG--NGR 678
QY 68 ---KGHNGALN--LNFNGKSAQNRFLTLTGTTNLNGKI 99
Db 679 YSLKVAQAARNTARLNLRQPNKVLVPGYVENGKI 715

Search completed: May 1, 2001, 15:03:04
Job time: 320 sec

THIS PAGE BLANK (USPTO)

Db 61 FFDEENGKGHNGALNLFNGKSAQNRFLLTGGANLNG 104
|||||

RESULT 2
US-09-142-970-5
; Sequence 5, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtmann
; APPLICANT: Monique Moreau
; TITLE OF INVENTION: IGAL PROTEASE FRAGMENT AS CARRIER
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE: 7101/0E616USO
; CURRENT APPLICATION NUMBER: US/09/142,970
; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-09-142-970-5

Query Match 92.1%; Score 524; DB 15; Length 104;
Best Local Similarity 97.9%; Pred. No. 3.9e-56;
Matches 95; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 1 LYYKNRYALKSGGRLNAPMPENGVAENNDWVFMGYTQEEARKNAMNNRRIGDFFG 60
|||||
Db 1 LYYKNRYALKSGGRLNAPMPENGVAENNDWVFMGYTQEEARKNAMNNRRIGDFFG 60
Qy 61 FFDEENGKGHNGALNLFNGKSAQNRFLLTGGANLNG 97
|||||
Db 61 FFDEENGKGHNGALNLFNGKSAQNRFLLTGGANLNG 97

RESULT 3
US-09-142-970-2
; Sequence 2, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtmann
; APPLICANT: Monique Moreau
; TITLE OF INVENTION: IGAL PROTEASE FRAGMENT AS CARRIER
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE: 7101/0E616USO
; CURRENT APPLICATION NUMBER: US/09/142,970
; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-142-970-2

Query Match 83.5%; Score 475; DB 15; Length 104;
Best Local Similarity 88.7%; Pred. No. 3.8e-50;
Matches 86; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
Qy 1 LYYKNRYALKSGGRLNAPMPENGVAENNDWVFMGYTQEEARKNAMNNRRIGDFFG 60
|||||
Db 1 LYYKNRYALKSGGRLNAPMPENGVAENNDWVFMGYTQEEARKNAMNNRRIGDFFG 60

Qy 61 FFDEENGKGHNGALNLFNGKSAQNRFLLTGGANLNG 97
|||||
Db 61 FFDEENGKGHNGALNLFNGKSAQNRFLLTGGANLNG 97

RESULT 4
US-09-142-970-3
; Sequence 3, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtmann
; APPLICANT: Monique Moreau
; TITLE OF INVENTION: IGAL PROTEASE FRAGMENT AS CARRIER
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE: 7101/0E616USO
; CURRENT APPLICATION NUMBER: US/09/142,970
; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-142-970-3

Query Match 81.2%; Score 462; DB 15; Length 104;
Best Local Similarity 86.6%; Pred. No. 1.5e-48;
Matches 84; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
Qy 1 LYYKNRYALKSGGRLNAPMPENGVAENNDWVFMGYTQEEARKNAMNNRRIGDFFG 60
|||||
Db 1 LYYKNRYALKSGGRLNAPMPENGVAENNDWVFMGYTQEEARKNAMNNRRIGDFFG 60
Qy 61 FFDEENGKGHNGALNLFNGKSAQNRFLLTGGANLNG 97
|||||
Db 61 FFDEENGKGHNGALNLFNGKSAQNRFLLTGGANLNG 97

RESULT 5
US-08-578-788-4
; Sequence 4, Application US/08578788
; GENERAL INFORMATION:
; APPLICANT: KILIAN, Mogens
; APPLICANT: POULSEN, Knud
; TITLE OF INVENTION: IMMUNOGLOBULIN A1 PROTEASES (IGAL
; TITLE OF INVENTION: PROTEASES), METHOD OF INDUCTION OF BROAD-SPECTRUM
; TITLE OF INVENTION: INHIBITORY ANTIBODIES USING GENE-TECHNOLOGICALLY PRODUCED
; TITLE OF INVENTION: BACTERIAL IGAL PROTEASES APPLIED ALONE OR CONJUGATED TO BA
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN
; STREET: 400 Seventh Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578,788
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,638
; FILING DATE: 22-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: PRICE, D. DOUGLAS

; REGISTRATION NUMBER: 24,514
; REFERENCE/DOCKET NUMBER: 162/P54307A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-6666
; TELEFAX: (202) 393-5350
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1560 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-578-788-4

Query Match 80.5%; Score 458; DB 9; Length 1560;
Best Local Similarity 77.7%; Pred. No. 1.6e-46;
Matches 87; Conservative 5; Mismatches 10; Indels 10; Gaps 2;

Qy 1 LYXKNRYALKSGRLNAPMPENGVAENNDVPMGYTQEEARKNAMNNKRRIGDFGG 60
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 584 LYXKNRYALKSGSVNAPMPENGQTNNDWILMGSTQEEAKKNAMNNKRRIGDFGG 643
Qy 61 FFDENGKGHNALNPNFGKSAQNRELLTGGANLNG-----GN---GRP 102
|| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 644 FFDENGKGHNALNPNFGKSAQNRELLTGGTNLNGKISVTQGNVLLSGRP 695

RESULT 6

US-09-142-970-1

; Sequence 1, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtman
; TITLE OF INVENTION: IGAL PROTEASE FRAGMENT AS CARRIER
; FILE REFERENCE: 7101/OE616050
; CURRENT APPLICATION NUMBER: US/09/142,970
; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-142-970-1

Query Match 79.6%; Score 453; DB 15; Length 104;
Best Local Similarity 84.5%; Pred. No. 1.8e-47;
Matches 82; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LYXKNRYALKSGRLNAPMPENGVAENNDVPMGYTQEEARKNAMNNKRRIGDFGG 60
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 1 LYXKNRYALKSGSVNAPMPENGQTNNDWILMGSTQEEAKKNAMNNKRRIGDFGG 60
Qy 61 FFDENGKGHNALNPNFGKSAQNRELLTGGANLNG 97
|| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 61 FFDENGKGHNALNPNFGKSAQNRELLTGGTNLNG 97

RESULT 7

US-09-155-077-1

; Sequence 1, Application US/09155077
; GENERAL INFORMATION:
; APPLICANT: Moreau, Monique
; APPLICANT: Mistretta, Noelle
; TITLE OF INVENTION: Polysaccharide-Peptide-Conjugates
; FILE REFERENCE: MBH99-777
; CURRENT APPLICATION NUMBER: US/09/155,077

; CURRENT FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Artificial/Unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(105)
; OTHER INFORMATION: Synthetic 105 mer peptide
US-09-155-077-1

Query Match 79.6%; Score 453; DB 15; Length 105;
Best Local Similarity 84.5%; Pred. No. 1.9e-47;
Matches 82; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LYXKNRYALKSGRLNAPMPENGVAENNDVPMGYTQEEARKNAMNNKRRIGDFGG 60
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 2 LYXKNRYALKSGSVNAPMPENGQTNNDWILMGSTQEEAKKNAMNNKRRIGDFGG 61
Qy 61 FFDENGKGHNALNPNFGKSAQNRELLTGGANLNG 97
|| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 62 FFDENGKGHNALNPNFGKSAQNRELLTGGTNLNG 98

RESULT 8

US-08-578-788-2

; Sequence 2, Application US/08578788
; GENERAL INFORMATION:
; APPLICANT: KILIAN, Mogens
; APPLICANT: POULSEN, Knud
; TITLE OF INVENTION: IMMUNOGLOBULIN A1 PROTEASES (IGAL
; TITLE OF INVENTION: PROTEASES), METHOD OF INDUCTION OF BROAD-SPECTRUM
; TITLE OF INVENTION: INHIBITORY ANTIBODIES USING GENE-TECHNOLOGICALLY PRODUCED
; TITLE OF INVENTION: BACTERIAL IGAL PROTEASES APPLIED ALONE OR CONJUGATED TO BA
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN
; STREET: 400 Seventh Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578,788
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,638
; FILING DATE: 22-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: PRICE, D. DOUGLAS
; REGISTRATION NUMBER: 24,514
; REFERENCE/DOCKET NUMBER: 162/P54307A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-6666
; TELEFAX: (202) 393-5350
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1540 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-578-788-2

Query Match

49.5%; Score 281.5; DB 9; Length 1540;

```
Best Local Similarity 48.2%; Pred. No. 5.7e-25;
Matches 54; Conservative 17; Mismatches 30; Indels 11; Gaps 2;

QY 1 LYKKNRYALKSGRLNAPMPENGVAENNDWVFVGYTQEEARKNMNNKNNRRIGDFGG 60
   :|||:||||:| :||:| :||:||||:| :||:| :||:| :||:| :||:| :||:|
Db 595 LNLENTYYALKRGASTRSELPKNGSGESNEWLYMGKTSDEAKRNMNNHNNRNGFNG 654
QY 61 PFDEENGKNGHGNALNLFNGKSAQNRFLLTGANLNGG-----NGRP 102
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 655 YFGEPEGK-NNGNLNVTFKGSQNRFLLTGTGTLNGLDITVEKGTFLSLGRP 705

RESULT 9
US-08-296-791-3
; Sequence 3, Application US/08296791
; GENERAL INFORMATION:
; APPLICANT: St. Geme III, Joseph W.
; APPLICANT: Falkow, Stanley
; TITLE OF INVENTION: Haemophilus Adherence and Penetration
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1541 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
US-08-296-791-3

Query Match 49.5%; Score 281.5; DB 6; Length 1541;
Best Local Similarity 48.2%; Pred. No. 5.7e-25;
Matches 54; Conservative 17; Mismatches 30; Indels 11; Gaps 2;

QY 1 LYKKNRYALKSGRLNAPMPENGVAENNDWVFVGYTQEEARKNMNNKNNRRIGDFGG 60
   :|||:||||:| :||:| :||:||||:| :||:| :||:| :||:| :||:| :||:|
Db 595 LNLENTYYALKRGASTRSELPKNGSGESNEWLYMGKTSDEAKRNMNNHNNRNGFNG 654
QY 61 PFDEENGKNGHGNALNLFNGKSAQNRFLLTGANLNGG-----NGRP 102
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 655 YFGEPEGK-NNGNLNVTFKGSQNRFLLTGTGTLNGLDITVEKGTFLSLGRP 705

RESULT 10
US-08-296-791-5
; Sequence 5, Application US/08296791
; GENERAL INFORMATION:
; APPLICANT: St. Geme III, Joseph W.
; APPLICANT: Falkow, Stanley
; TITLE OF INVENTION: Haemophilus Adherence and Penetration
```

```
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1702 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
US-08-296-791-5

Query Match 48.4%; Score 275.5; DB 6; Length 1702;
Best Local Similarity 47.3%; Pred. No. 3.5e-24;
Matches 53; Conservative 17; Mismatches 31; Indels 11; Gaps 2;

QY 1 LYKKNRYALKSGRLNAPMPENGVAENNDWVFVGYTQEEARKNMNNKNNRRIGDFGG 60
   :|||:||||:| :||:| :||:||||:| :||:| :||:| :||:| :||:| :||:|
Db 601 LNLENTYYALKRGASTRSELPKNGSGESNEWLYMGKTSDAKRVNMNNHNNRNGFNG 660
QY 61 PFDEENGKNGHGNALNLFNGKSAQNRFLLTGANLNGG-----NGRP 102
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 661 YFGEPEGK-NNGNLNVTFKGSQNRFLLTGTGTLNGLDITVEKGTFLSLGRP 711

RESULT 11
US-08-296-791-4
; Sequence 4, Application US/08296791
; GENERAL INFORMATION:
; APPLICANT: St. Geme III, Joseph W.
; APPLICANT: Falkow, Stanley
; TITLE OF INVENTION: Haemophilus Adherence and Penetration
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

REFLECTANT: Liu, Qingdong
; APPLICANT: Lutfiyva, Linda L.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2001, 15:03:04 ; Search time 11.48 Seconds
(without alignments)
37.203 Million cell u

Title: US-09-142-970-4

Perfect score: 569
Sequence: 1 LYYNRYRYALKSGGRLNAP.....NRFLTGGANLNGNGRPPVK 104

Scoring table: BLOSUM62

decreasing cubic: DECSOM2
 : Gapop 10.0 , Gapext 0.5

Searched: 25343 seqs, 4106610 residues

Total number of hits satisfying chosen parameters: 25343

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New:*

1: /cqn2_6/ptodata1/paa/pct_NEW_COMB.pcp.*
2: /cqn2_6/ptodata1/paa/US06_NEW_COMB.pcp.*
3: /cqn2_6/ptodata1/paa/US07_NEW_COMB.pcp.*
4: /cqn2_6/ptodata1/paa/US08_NEW_COMB.pcp.*
5: /cqn2_6/ptodata1/paa/US09_NEW_COMB.pcp.*
6: /cqn2_6/ptodata1/paa/US60_NEW_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	63	11.1	1455	6	US-60-248-823-105	Sequence 105, Appl
2	62.5	11.0	421	5	US-09-420-785A-4	Sequence 4, Appl
3	61	10.7	302	5	US-09-308-823A-216	Sequence 216, Appl
4	61	10.7	342	5	US-09-509-031-6	Sequence 6, Appl
5	61	10.7	495	5	US-09-509-031-4	Sequence 4, Appl
6	58.5	10.3	137	5	US-09-739-449-8856	Sequence 8856, Appl
7	57.5	10.1	409	5	US-09-792-024-92	Sequence 92, Appl
8	57.5	10.1	495	5	US-09-819-386-5	Sequence 5, Appl
9	57.5	10.1	498	5	US-09-819-386-2	Sequence 2, Appl
10	57	10.0	417	5	US-09-392-846-2	Sequence 2, Appl
11	57	10.0	730	6	US-60-248-505-921	Sequence 921, Appl
12	56.5	9.9	1008	5	US-09-308-453-2	Sequence 2, Appl
13	55.5	9.8	143	5	US-09-485-558A-19	Sequence 19, Appl
14	55	9.7	482	5	US-09-509-031-16	Sequence 16, Appl
15	55	9.7	597	5	US-09-739-449-10335	Sequence 10335, A
16	54.5	9.6	397	5	US-09-463-238-15	Sequence 15, Appl
17	54.5	9.6	766	5	US-09-463-238-6	Sequence 6, Appl
18	54	9.5	461	5	US-09-739-449-11869	Sequence 11869, A
19	53.5	9.4	409	5	US-09-813-329-2	Sequence 2, Appl
20	53.5	9.4	541	5	US-09-739-449-8767	Sequence 8767, Ap
21	53.5	9.4	580	5	US-09-739-449-8757	Sequence 8757, Ap
22	53.5	9.4	884	6	US-60-248-505-690	Sequence 690, App
23	53.5	9.4	1355	6	US-60-248-505-1205	Sequence 1205, Ap
24	52.5	9.2	220	5	US-09-739-449-11308	Sequence 11308, A
25	52.5	9.2	475	5	US-09-772-393-2	Sequence 2, Appl
26	52.5	9.2	610	5	US-09-266-091-2	Sequence 2, Appl
27	52.5	9.2	610	5	US-09-802-640-36	Sequence 36, Appl

28	52.5	9.2	793	5	US-09-463-238-5	Sequence 5, Appli
29	52.5	9.2	838	5	US-09-463-238-14	Sequence 14, Appli
30	52.5	9.2	940	5	US-09-818-879-4	Sequence 4, Appli
31	52	9.1	1096	5	US-09-792-024-120	Sequence 120, App
32	51.5	9.1	392	5	US-09-739-449-9678	Sequence 9678, App
33	51.5	9.1	793	5	US-09-739-449-11850	Sequence 11850, A
34	51	9.0	238	5	US-09-739-449-11393	Sequence 11393, A
35	51	9.0	390	5	US-09-739-449-11991	Sequence 11991, A
36	51	9.0	554	5	US-09-813-408-28	Sequence 28, Appli
37	51	9.0	791	5	US-09-804-472-2	Sequence 2, Appli
38	50	8.8	249	5	US-09-739-449-8712	Sequence 8712, Ap
39	50	8.8	266	5	US-09-792-024-72	Sequence 72, Appli
40	50	8.8	391	5	US-09-813-339-9	Sequence 9, Appli
41	50	8.8	417	5	US-09-514-650A-13	Sequence 13, Appli
42	50	8.8	617	6	US-60-248-505-741	Sequence 741, App
43	50	8.8	662	5	US-09-739-449-11810	Sequence 11810, A
44	50	8.8	684	5	US-09-823-240-9	Sequence 9, Appli
45	50	8.8	1032	5	US-09-739-449-11325	Sequence 11325, A

ALIGNMENTS

```

RESULT      1
US-60-248-823-105
; Sequence 105, Application US/60248823
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN LIPASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN LIPASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000949
; CURRENT APPLICATION NUMBER: US/60/248,823
; CURRENT FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 189
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 105
; LENGTH: 1455
; TYPE: PRT
; ORGANISM: HUMAN
; US-60-248-823-105

```

```

Query Match      11.1%; Score 63; DB 6; Length 1455;
Best Local Similarity 22.7%; Pred. No. 8.1;
Matches 15; Conservative 14; Mismatches 31; Indels 6; Gaps 1;

QY 16 RLNAPMPENGVAENNDVFMGYCTOEEARKKNAMNNKRRIRGDFCGFPDEENGKCHNGALN 75
   || :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:| :|:|
Db 770 RLSEVPQTNHAHESKEWYHASLTRAQAEHMLM-----RVPRDGAFLVRRKRNPEPNSYAIS 823

QY 76 LNFNGK 81
   ||
Db 824 FRAEGK 829
   ||

```


Db 58 YRNFKDANIKOLGRNAA-----GIAQNG-----CNLG--F 87
QY 63 DEENGKGHNGALNLFNGKS-----AONRELLTGGANL--NGNG 100
Db 88 IROGRNGHSATLQONGNANNAYGFOYGRNT-----GTNVQDGDNG 128

RESULT 7
US-09-792-024-92
; Sequence 92, Application US/09792024
; GENERAL INFORMATION:
; APPLICANT: Roemer, Terry
; APPLICANT: Jiang, Bo
; APPLICANT: Boone, Charles
; APPLICANT: Bussey, Howard
; TITLE OF INVENTION: Gene Disruption Methodologies for Drug
; FILE REFERENCE: 10182-004-999
; CURRENT APPLICATION NUMBER: US/09/792,024
; CURRENT FILING DATE: 2001-02-20
; NUMBER OF SEQ ID NOS: 490
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 92
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-792-024-92

Query Match 10.1%; Score 57.5; DB 5; Length 409;
Best Local Similarity 31.1%; Pred. No. 6.9;
Matches 14; Conservative 7; Mismatches 21; Indels 3; Gaps 1;
QY 8 YYALKSGRLNAPMPENGVAENNDWVFVGYTQEEARKNMMNKN 52
Db 101 YKLLSKNGKFGDLP---LDQNLVDLTGTFTIEEEKEKEENN 142

RESULT 8
US-09-819-386-5
; Sequence 5, Application US/09819386
; GENERAL INFORMATION:
; APPLICANT: THORPE, PHILIP E.
; APPLICANT: RAN, SOPHIA
; TITLE OF INVENTION: CANCER TREATMENT METHODS USING THERAPEUTIC CONJUGATES
; FILE REFERENCE: 4001.002300
; CURRENT APPLICATION NUMBER: US/09/819,386
; CURRENT FILING DATE: 2001-03-28
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-819-386-5

Query Match 10.1%; Score 57.5; DB 5; Length 495;
Best Local Similarity 21.7%; Pred. No. 8.9;
Matches 28; Conservative 20; Mismatches 42; Indels 39; Gaps 7;
QY 2 YYKNRYALKSGRLNAPMPENGVAENNDWVFVGYTQEE--ARKNAMNNKNNRRIGDFG 59
Db 334 FQGWKEYKMGFGN-----PSGEYWLGNFEFAITSQRYMLRIELMDWEGNRAYSQYD 387
QY 60 GFF---DEENG---KGNGAL-----NLNFKGSAQN-----RFLLTGG--- 92
Db 388 RFHIGNEKQNYRLYKGTGTACKQSSLLIHGADFSTKDADNDNCMKCALMLTGGWFFD 447
QY 93 ----ANLNG 97

Db 448 ACGPSNLNG 456

RESULT 9
US-09-819-386-2
; Sequence 2, Application US/09819386
; GENERAL INFORMATION:
; APPLICANT: THORPE, PHILIP E.
; APPLICANT: RAN, SOPHIA
; TITLE OF INVENTION: CANCER TREATMENT METHODS USING THERAPEUTIC CONJUGATES
; FILE REFERENCE: 4001.002300
; CURRENT APPLICATION NUMBER: US/09/819,386
; CURRENT FILING DATE: 2001-03-28
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-819-386-2

Query Match 10.1%; Score 57.5; DB 5; Length 498;
Best Local Similarity 21.7%; Pred. No. 9;
Matches 28; Conservative 20; Mismatches 42; Indels 39; Gaps 7;
QY 2 YYKNRYALKSGRLNAPMPENGVAENNDWVFVGYTQEE--ARKNAMNNKNNRRIGDFG 59
Db 337 FQGWKEYKMGFGN-----PSGEYWLGNFEFAITSQRYMLRIELMDWEGNRAYSQYD 390
QY 60 GFF---DEENG---KGNGAL-----NLNFKGSAQN-----RFLLTGG--- 92
Db 391 RFHIGNEKQNYRLYKGTGTACKQSSLLIHGADFSTKDADNDNCMKCALMLTGGWFFD 450
QY 93 ----ANLNG 97
Db 451 ACGPSNLNG 459

RESULT 10
US-09-392-846-2
; Sequence 2, Application US/09392846
; GENERAL INFORMATION:
; APPLICANT: Pienkos, Philip T.
; APPLICANT: Kerridge, Alison
; APPLICANT: Holland, Herbert L.
; TITLE OF INVENTION: Method of Producing Chiral Sulfoxides
; FILE REFERENCE: 1405.1103-000
; CURRENT APPLICATION NUMBER: US/09/392,846
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Rhodococcus species
US-09-392-846-2

Query Match 10.0%; Score 57; DB 5; Length 417;
Best Local Similarity 30.0%; Pred. No. 8.2;
Matches 12; Conservative 9; Mismatches 17; Indels 2; Gaps 1;
QY 12 KSGRLNAPMPEN--GVAENNDWVFVGYTQEEARKNAMNN 49
Db 183 QQGAIIAAIPTSRAGVTPNDNAAIGNROTDSRSTDFHN 222
RESULT 11

```
US-60-248-505-921
; FILE REFERENCE: BB-1107
; CURRENT APPLICATION NUMBER: US/09/485,558A
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USES THEREOF
; FILE REFERENCE: C1000918
; CURRENT APPLICATION NUMBER: US/60/248,505
; CURRENT FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 1998
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 921
; LENGTH: 730
; TYPE: PRT
; ORGANISM: Human
US-60-248-505-921

Query Match          10.0%; Score 57; DB 6; Length 730;
Best Local Similarity 24.7%; Pred. No. 17;
Matches 18; Conservative 9; Mismatches 20; Indels 26; Gaps 2;

QY 11 LKSGRLNAPMPENGVAENNDWVFMGYTQ---EARKNAM----- 47
DB 642 VRENDRLRVKVEERGVAENSMVISNGDTNLSLRKAKKRAFOLEDEETEPDYKYSKKHK 701
QY 48 ---NNKNNRRIGD 57
DB 702 ROENNNNEKVF 714

RESULT 12
US-09-308-453-2
; Sequence 2, Application US/09308453
; GENERAL INFORMATION:
; APPLICANT: Roche Diagnostics GMBH
; TITLE OF INVENTION: Recombinant collagenase type I from Clostridium histolyticum and
; TITLE OF INVENTION: for isolating cells and groups of cells
; FILE REFERENCE: BMID9924US
; CURRENT APPLICATION NUMBER: US/09/308,453
; CURRENT FILING DATE: 1999-09-09
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 1008
; TYPE: PRT
; ORGANISM: Clostridium histolyticum
US-09-308-453-2

Query Match          9.9%; Score 56.5; DB 5; Length 1008;
Best Local Similarity 21.2%; Pred. No. 30;
Matches 17; Conservative 19; Mismatches 29; Indels 15; Gaps 3;

QY 8 YYALSGGRLNAPMPENGVAENNDWVFMGYTQEEARKNAMNNKNNRRIGDFGFFDEENG 67
DB 814 YFDVVEDGDVTIELPYSG-SSNFTWLTVYKEGDQNHIAASGIDKNNKVGTF-----KAT 866
QY 68 KG-----HNGALNLFN 79
DB 867 KGRHYVFIYKHDASASNISYS 886

RESULT 13
US-09-485-558A-19
; Sequence 19, Application US/09485558A
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Liu, Zhan-Bin
; APPLICANT: E. I. du Pont de Nemours and Company
; TITLE OF INVENTION: Plant Genes Encoding Drl and DraPI, a Global Repressor
; TITLE OF INVENTION: Complex of Transcription
```

```
; FILE REFERENCE: BB-1107
; CURRENT APPLICATION NUMBER: US/09/485,558A
; GENERAL INFORMATION:
; APPLICANT: Microsoft Word Version 7.0A
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 19
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Glycine max
US-09-485-558A-19

Query Match          9.8%; Score 55.5; DB 5; Length 143;
Best Local Similarity 27.0%; Pred. No. 3;
Matches 20; Conservative 8; Mismatches 35; Indels 11; Gaps 2;

QY 25 GVAENNDWVFMGYTQEEARKNAMNNKNNRRIGDFGFFDEENGKNGHNGALNLFNPKNSAQ 84
DB 66 GFCEYTEEYVAAEQHK-----LETMQDSLGGGG-----GKWNNGAEMTEEALABQ 114
QY 85 NRFLLTGGANLNGG 98
DB 115 ORMLAEARARMNGG 128

RESULT 14
US-09-509-031-16
; Sequence 16, Application US/09509031
; GENERAL INFORMATION:
; APPLICANT: Koentgen, Frank
; APPLICANT: Suess, Gabriele M.
; APPLICANT: Tarlinton, David M.
; APPLICANT: Treutlein, Herbert R.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES AND A METHOD OF PRODUCING SAME
; FILE REFERENCE: 13474
; CURRENT APPLICATION NUMBER: US/09/509,031
; CURRENT FILING DATE: 2000-06-09
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: ccMTLgI protein
; OTHER INFORMATION: sequence
US-09-509-031-16

Query Match          9.7%; Score 55; DB 5; Length 482;
Best Local Similarity 25.4%; Pred. No. 17;
Matches 29; Conservative 14; Mismatches 35; Indels 36; Gaps 7;

QY 14 GGLNAPMPENGVAENNDWVFM---GYTOEEA--RKNAMNNKNNRRIGDFGFFD----- 63
DB 296 GGGSGAGGGGGGGSEN--LYFGQGGGSAEYTIKANLIFANGSTQTAEFKGTPEKATSE 353
QY 64 -----EENG-----KGHNGALNLFNFGSAQNRELLTGANLNGNG 100
DB 354 AYAYADTLKDNGEYTVVDVADKGY--TLNKFAGKEASG-----GGSGGGGGSG 400

RESULT 15
US-09-739-449-10335
; Sequence 10335, Application US/09739449
; GENERAL INFORMATION:
; APPLICANT: Hinkie, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)C
; CURRENT APPLICATION NUMBER: US/09/739,449
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 09/514,000
```


Search completed: May 1, 2001, 15:03:05
Job time: 321 sec

THIS PAGE BLANK (USPTO)

Qy	1	LYXKRYRYALKSGRLNAPENGVAENNDWIFMGYTOEARKNAAMHNRRIGDGG	60
Db	1	LYXKRYRYALKSGRLNAPENGVAENNDWIFMGYTOEARKNAAMHNRRIGDGG	60
Qy	61	FDEENGRGHGALNLFNGSKAONRELLTGGANLNGKISVTQG	104

Db 61 PFDEENGKNGHNGALNLFNGKSAQNRFLLTGGANLNGKISVTQG 104
|||||

RESULT 2
US-09-142-970-4

; Sequence 4, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtman
; APPLICANT: Monique Moreau
; TITLE OF INVENTION: ICAL PROTEASE FRAGMENT AS CARRIER
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE: 7101/OE616USO
; CURRENT APPLICATION NUMBER: US/09/142,970
; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-09-142-970-4

Query Match 92.7%; Score 524; DB 15; Length 104;
Best Local Similarity 97.9%; Pred. No. 6.3e-58;
Matches 95; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LYKKNRYALKSGGRLNAPMPENGVAENNDWIFMGYTQEEARKNAMHKNRRIGDFGG 60
|||||

Db 1 LYKKNRYALKSGGRLNAPMPENGVAENNDWIFMGYTQEEARKNAMHKNRRIGDFGG 60
|||||

Qy 61 PFDEENGKNGHNGALNLFNGKSAQNRFLLTGGANLNG 97
|||||

Db 61 PFDEENGKNGHNGALNLFNGKSAQNRFLLTGGANLNG 97
|||||

RESULT 3
US-09-142-970-2

; Sequence 2, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtman
; APPLICANT: Monique Moreau
; TITLE OF INVENTION: ICAL PROTEASE FRAGMENT AS CARRIER
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE: 7101/OE616USO
; CURRENT APPLICATION NUMBER: US/09/142,970
; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-142-970-2

Query Match 91.0%; Score 514; DB 15; Length 104;
Best Local Similarity 89.4%; Pred. No. 1.1e-56;
Matches 93; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Qy 1 LYKKNRYALKSGGRLNAPMPENGVAENNDWIFMGYTQEEARKNAMHKNRRIGDFGG 60
|||||

Db 1 LYKKNRYALKSGGRLNAPMPENGVAENNDWIFMGYTQEEARKNAMHKNRRIGDFGG 60
|||||

Qy 61 PFDEENGKNGHNGALNLFNGKSAQNRFLLTGGANLNGKISVTQG 104
|||||

Db 61 PFDEENGKNGHNGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 104
|||||

RESULT 4

US-09-142-970-3
; Sequence 3, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtman
; APPLICANT: Monique Moreau
; TITLE OF INVENTION: ICAL PROTEASE FRAGMENT AS CARRIER
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE: 7101/OE616USO
; CURRENT APPLICATION NUMBER: US/09/142,970
; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-142-970-3

Query Match 88.7%; Score 501; DB 15; Length 104;
Best Local Similarity 87.5%; Pred. No. 4.9e-55;
Matches 91; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

Qy 1 LYKKNRYALKSGGRLNAPMPENGVAENNDWIFMGYTQEEARKNAMHKNRRIGDFGG 60
|||||

Db 1 LYKKNRYALKSGGRLNAPMPENGQTEENNDWIFMGYTQEEARKNAMHKNRRIGDFGG 60
|||||

Qy 61 PFDEENGKNGHNGALNLFNGKSAQNRFLLTGGANLNGKISVTQG 104
|||||

Db 61 PFDEENGKNGHNGALNLFNGKSAQNRFLLTGGTNLNGKISVTQG 104
|||||

RESULT 5

US-09-142-970-1
; Sequence 1, Application US/09142970
; GENERAL INFORMATION:
; APPLICANT: Mark Achtman
; APPLICANT: Monique Moreau
; TITLE OF INVENTION: ICAL PROTEASE FRAGMENT AS CARRIER
; TITLE OF INVENTION: PEPTIDE
; FILE REFERENCE: 7101/OE616USO
; CURRENT APPLICATION NUMBER: US/09/142,970
; CURRENT FILING DATE: 1999-04-02
; EARLIER APPLICATION NUMBER: EP 97100883.4
; EARLIER FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: PCT/EP98/00294
; EARLIER FILING DATE: 1998-01-20
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
US-09-142-970-1

Query Match 87.4%; Score 494; DB 15; Length 104;
Best Local Similarity 87.5%; Pred. No. 3.7e-54;
Matches 91; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

Qy 1 LYKKNRYALKSGGRLNAPMPENGVAENNDWIFMGYTQEEARKNAMHKNRRIGDFGG 60
|||||

Db 1 LYKKNRYALKSGGRLNAPMPENGQTEENNDWILMGSTQEEARKNAMHKNRRIGDFGG 60
|||||

QY 61 FFDENGKGHGALNLFNGKSAQNRFLLTGGANLNGKISVTQG 104
|||
Db 61 FFGEENGKGHGALNLFNGKSAQNRFLLTGGANLNGKISVTQG 104

RESULT 6

US-09-155-077-1
; Sequence 1, Application US/09155077
; GENERAL INFORMATION:
; APPLICANT: Moreau, Monique
; APPLICANT: Mistretta, Noelle
; TITLE OF INVENTION: Polysaccharide-Peptide-Conjugates
; FILE REFERENCE: MBH99-777
; CURRENT APPLICATION NUMBER: US/09/155,077
; CURRENT FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Artificial/Unknown
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(105)
; OTHER INFORMATION: Synthetic 105 mer peptide
US-09-155-077-1

Query Match 87.4%; Score 494; DB 15; Length 105;
Best Local Similarity 87.5%; Pred. No. 3.7e-54;
Matches 91; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 LYKKNRYVALKSGGRNLNAPMPENGVAENNDWIFMGYTQEEAKKNAMHNKNNRRIGDFGG 60
|||
Db 2 LYKKNRYVALKSGGSVNAPMPENGQTNNDWILMGSTQEEAKKNAMHNKNNRRIGSFG 61
|||
QY 61 FFDENGKGHGALNLFNGKSAQNRFLLTGGANLNGKISVTQG 104
|||
Db 62 FFGEENGKGHGALNLFNGKSAQNRFLLTGGANLNGKISVTQG 105
|||

RESULT 7

US-08-578-788-4
; Sequence 4, Application US/08578788
; GENERAL INFORMATION:
; APPLICANT: KILIAN, Mogens
; APPLICANT: POULSEN, Knud
; TITLE OF INVENTION: IMMUNOGLOBULIN A1 PROTEASES (IGAL
; TITLE OF INVENTION: PROTEASES), METHOD OF INDUCTION OF BROAD-SPECTRUM
; TITLE OF INVENTION: INHIBITORY ANTIBODIES USING GENE-TECHNOLOGICALLY PRODUCED
; TITLE OF INVENTION: BACTERIAL IGAL PROTEASES APPLIED ALONE OR CONJUGATED TO BA
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN
; STREET: 400 Seventh Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578,788
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,638
; FILING DATE: 22-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: PRICE, D. DOUGLAS
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578,788
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,638
; FILING DATE: 22-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: PRICE, D. DOUGLAS

; REGISTRATION NUMBER: 24,514
; REFERENCE/DOCKET NUMBER: 162/P54307A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-6666
; TELEFAX: (202) 393-5350
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1560 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-578-788-4

Query Match 87.4%; Score 494; DB 9; Length 1560;
Best Local Similarity 87.5%; Pred. No. 1.4e-52;
Matches 91; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 1 LYKKNRYVALKSGGRNLNAPMPENGVAENNDWIFMGYTQEEAKKNAMHNKNNRRIGDFGG 60
|||
Db 584 LYKKNRYVALKSGGSVNAPMPENGQTNNDWILMGSTQEEAKKNAMHNKNNRRIGSFG 643
|||
QY 61 FFDENGKGHGALNLFNGKSAQNRFLLTGGANLNGKISVTQG 104
|||
Db 644 FFGEENGKGHGALNLFNGKSAQNRFLLTGGANLNGKISVTQG 687
|||

RESULT 8

US-08-578-788-2
; Sequence 2, Application US/08578788
; GENERAL INFORMATION:
; APPLICANT: KILIAN, Mogens
; APPLICANT: POULSEN, Knud
; TITLE OF INVENTION: IMMUNOGLOBULIN A1 PROTEASES (IGAL
; TITLE OF INVENTION: PROTEASES), METHOD OF INDUCTION OF BROAD-SPECTRUM
; TITLE OF INVENTION: INHIBITORY ANTIBODIES USING GENE-TECHNOLOGICALLY PRODUCED
; TITLE OF INVENTION: BACTERIAL IGAL PROTEASES APPLIED ALONE OR CONJUGATED TO BA
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: JACOBSON, PRICE, HOLMAN & STERN
; STREET: 400 Seventh Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/578,788
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/215,638
; FILING DATE: 22-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: PRICE, D. DOUGLAS
; REGISTRATION NUMBER: 24,514
; REFERENCE/DOCKET NUMBER: 162/P54307A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-6666
; TELEFAX: (202) 393-5350
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1540 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-578-788-2

Query Match 52.8%; Score 298.5; DB 9; Length 1540;

```

; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: A-59941/RFT/RMS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1545 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; US-08-296-791-4

Query Match 51.6%; Score 291.5; DB 6; Length 11
Best Local Similarity 52.8%; Pred. No. 4e-27;
Matches 56; Conservative 16; Mismatches 31; Indels

QY 1 LYY-KNYRYVYALKSGRGNAPWENGVAENNDWIFMGVTOEARKNAAMHNKH
DB 597 LYNEENRYVYALKKDASTRSEFPNGRGSNNWLYMGTEKADAKNAAMHNH
QY 59 GGFDEENGKHNGALNFNFGSKAQNRFLLTGGANLNGKISVTQG 104
DB 657 NGYFGEEGK-NNGNLNVTFKGSQNFLLTGGTNLNGDLNVQOG 701

RESULT 11
US-08-296-791-5
; Sequence 5, Application US/08296791
; GENERAL INFORMATION:
; APPLICANT: St. Geme III, Joseph W.
; APPLICANT: Falkow, Stanley
; TITLE OF INVENTION: Haemophilus Adherence and Penetration
; TITLE OF INVENTION: Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

```



```
; CURRENT APPLICATION NUMBER: US/09/733,089
; CURRENT FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 09/474,435
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: US 09/620,392
; PRIOR FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 24143
; SEQ ID NO 13711
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(450)
; OTHER INFORMATION: unsure at all xaa locations
US-09-733-089-13711

Query Match      15.3%  Score 86.5; DB 21; Length 450;
Best Local Similarity 24.0%  Pred. No. 0.045;
Matches 36; Conservative 19; Mismatches 40; Indels 55; Gaps 8;

QY 3 YKNRYYA-----LKSGRLN-----APM-----PENGVAENNDWIPIWG- 36
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3 YHIYIYSLVCTILMVAGYLNSTCXNFAPLYITVSFASIQRVSKGKITQKFWVDLGN 62
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 37 --YTOEEARKNAMNH-----KNRRIGDF-----GGFFDE-----ENCKGHN-- 71
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 63 LNLVQXXEKTIVHILPMSLLTVAHQSGGEIRFTESHGGGDDVSGDGEHCGCHDDG 122
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 72 -----GALNLFNGKSAQNRFLLTGTGANLNG 97
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 123 SCAGGSLGILNFNQVQGEVTFMQGSLVSG 152
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
US-09-270-767-41767
; Sequence 41767, Application US/09270767
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41767
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-41767

Query Match      13.3%  Score 75; DB 16; Length 419;
Best Local Similarity 24.1%  Pred. No. 1.1;
Matches 28; Conservative 10; Mismatches 32; Indels 46; Gaps 5;

QY 25 GVAENNDWIFMGYTQEEARKNAMHKNR-----RIGDFG-----GFF 62
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 274 GYGANN--MLGFNNLPGLMNSGNGFNNGGNGNFGNFGGNGGNGSGNF 330
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 63 DEENGK-----GHNGALNL-----NFGKSAQNRFLLTGTGANLNG 97
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 331 GNDGNSGNFNGFNNGGNGFNNGGNGGNGGNGFNNSPCGVNFGNNGGNSFGG 386
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: May 1, 2001, 15:02:52
Job time: 334 sec

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	63.5	11.2	421	5	US-09-420-785A-4	Sequence 4, Appli	
2	63	11.2	1455	6	US-60-248-823-105	Sequence 105, Appl	
3	61	10.8	342	5	US-09-509-031-6	Sequence 6, Appli	
4	61	10.8	495	5	US-09-509-031-4	Sequence 4, Appli	
5	60.5	10.7	495	5	US-09-819-386-5	Sequence 5, Appli	
6	60.5	10.7	498	5	US-09-819-386-2	Sequence 2, Appli	
7	58.5	10.4	409	5	US-09-792-024-92	Sequence 92, Appli	
8	58	10.3	137	5	US-09-793-449-8856	Sequence 8856, Ap	
9	56.5	10.0	1008	5	US-09-308-453-2	Sequence 2, Appli	
10	54.5	9.6	124	5	US-09-739-449-12040	Sequence 12040, A	
11	54.5	9.6	136	6	US-60-248-505-1051	Sequence 1051, A	
12	54	9.6	461	5	US-09-739-449-11869	Sequence 11869, A	
13	54	9.6	554	5	US-09-813-408-28	Sequence 28, Appl	
14	53.5	9.5	396	5	US-09-739-449-11819	Sequence 11819, A	
15	53	9.4	244	5	US-09-739-449-9411	Sequence 9411, Ap	
16	53	9.4	704	5	US-09-646-351-21	Sequence 21, Appl	
17	53	9.4	730	6	US-60-248-505-921	Sequence 921, App	
18	52.5	9.3	392	5	US-09-739-449-9678	Sequence 9678, Ap	
19	52.5	9.3	409	5	US-09-813-329-2	Sequence 2, Appli	
20	52.5	9.3	580	5	US-09-739-449-8757	Sequence 8757, Ap	
21	52.5	9.3	791	5	US-09-804-472-2	Sequence 2, Appli	
22	52.5	9.3	4563	5	US-09-802-640-32	Sequence 32, Appl	
23	52	9.2	281	5	US-09-792-024-118	Sequence 118, App	
24	52	9.2	302	5	US-09-308-823A-216	Sequence 216, App	
25	52	9.2	390	5	US-09-739-449-11991	Sequence 11991, A	
26	52	9.2	418	5	US-09-603-124B-32	Sequence 32, Appl	
27	52	9.2	940	5	US-09-818-879-4	Sequence 4, Appli	

```

RESULT 5
US-09-819-386-5
; Sequence 5, Application US/09819386
; GENERAL INFORMATION:
; APPLICANT: THORPE, PHILIP E.
; APPLICANT: PAN, SOPHIA
; TITLE OF INVENTION: CANCER TREATMENT METHODS USING THERAPEUTIC CONJUGATES THAT BIND TO AMINOPHOSPHOLIPIDS
; FILE REFERENCE: 4001.002300
; CURRENT APPLICATION NUMBER: US/09/819,386
; CURRENT FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US/09/351,457
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 5
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-819-386-5

```

		Query Match	10.7%; Score 60.5; DB 5; Length 495;
		Best Local Similarity	22.5%; Pred. No. 3.5;
Matches	29;	Conservative	19; Mismatches 42; Indels 39; Gaps
Qy	2	YKYNRYALKSGRLNAPMPENGVAENNDWIFMGYTQE--ARKNAMNHKNRRIGDFG	59
	:	:	:
Dd	334	FQRGWKEYKMFGFN-----PSEGYWLGNFEFFAITSORQYMLRIELMDWEGNRAYSQYD	387
	:	:	:
Qy	60	GFF-----DEENG-----KGHNAL-----NLNFNGKSQN-----RFLTGG-----	92
	:	:	:
Dd	388	RFHIGNEKQNYRLVYLGHTGTAGQSLLIHGHADFSTKDADNDNCMKCALMTLGGWFED	447
	:	:	:
Qy	93	----ANLNG	97
	:	:	:
Dd	448	ACGPSNLNG	456

```

Db      448  ACGPSNLNG  456      :||||
                                           :
RESULT      6
US-09-819-386-2
; Sequence 2, Application US/09819386
; GENERAL INFORMATION:
; APPLICANT: THORPE, PHILIP E.
; APPLICANT: RAN, SOPHIA
; TITLE OF INVENTION: CANCER TREATMENT METHODS USING THERAPEUTIC CONJUGATE
; TITLE OF INVENTION: THAT BIND TO AMINOPHOSPHOLIPIDS
; FILE REFERENCE: 4001.002300
; CURRENT APPLICATION NUMBER: US/09/819,386
; CURRENT FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US/09/351,457
; PRIOR FILING DATE: 1999-07-12
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Homo sapiens

```

RESULT 11

; TYPE: PRT
; ORGANISM: Agrobacterium tumefaciens
US-09-739-449-9411

Query Match 9.4%; Score 53; DB 5; Length 244;
Best Local Similarity 25.6%; Pred. No. 11;
Matches 22; Conservative 10; Mismatches 30; Indels 24; Gaps 5;
QY 8 YYALKSGG---RLN---APMPE--NGVAENNDWI--FMGYTQEEARKNAMNHK----- 50
| | : | | | | | : | | | : | | : | |
Db 120 YPVLDAGGGPIRLNPNNGGPIIVGLDGAIRQNDTIAASLGIFQADFSGFLRHPNSGVKPV 179
| | : | | | | | : | | | : | | : | |
QY 51 -----NNRRIGDFGGFFDEENGKG 69
| | : | | | | | : | | | : | | : | |
Db 180 AQPVPVNNHVEGVVQGYLEQSNVNG 205

Search completed: May 1, 2001, 15:03:05
Job time: 321 sec

THIS PAGE BLANK (USPTO)